

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 22:29:43 ; Search time 182.253 Seconds  
(without alignments)  
2597.404 Million cell updates/sec

Title: US-09-867-159A-1

Perfect score: 666

Sequence: 1 actaacgcctgcagatcaaa.....atccatatgtgtcattctc 666

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/ECTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	664.4	99.8	834	1	US-07-945-288-1
2	664.4	99.8	834	1	US-08-462-831-1
3	664.4	99.8	834	1	US-08-461-803-1
4	664.4	99.8	834	1	US-08-461-441-1
5	664.4	99.8	834	1	US-08-482-142-1
6	664.4	99.8	834	2	US-08-478-572-1
7	664.4	99.8	834	3	US-08-484-296-1
8	664.4	99.8	834	5	PCT-US93-08518-1
9	664.4	99.8	1172	1	US-07-945-288-9
10	664.4	99.8	1172	1	US-08-462-831-9
11	664.4	99.8	1172	1	US-08-461-441-9
12	664.4	99.8	1172	1	US-08-461-441-9
13	664.4	99.8	1172	5	PCT-US93-08518-9
14	635.2	95.4	857	3	US-08-460-040-1
15	469.8	70.5	1072	1	US-07-945-288-5
16	469.8	70.5	1072	1	US-08-462-831-5
17	469.8	70.5	1072	1	US-08-461-809-5
18	469.8	70.5	1072	1	US-08-461-441-5
19	469.8	70.5	1072	1	US-08-482-142-5
20	469.8	70.5	1072	2	US-08-478-572-5
21	469.8	70.5	1072	3	US-08-484-296-5
22	469.8	70.5	1072	5	PCT-US93-08518-5
23	150.4	22.6	549	1	US-08-482-142-194
24	150.4	22.6	549	2	US-08-478-572-194
25	150.4	22.6	549	3	US-08-484-296-194
26	90.4	13.6	549	1	US-08-482-142-192
27	90.4	13.6	549	2	US-08-478-572-192

28 90.4 13.6 549 3 US-08-484-296-192 Sequence 192, App  
29 90 13.5 549 1 US-08-482-142-196 Sequence 196, App  
30 90 13.5 549 2 US-08-478-572-196 Sequence 196, App  
31 90 13.5 549 3 US-08-484-296-196 Sequence 196, App  
32 56.4 8.5 2167 4 US-09-325-932A-102 Sequence 102, App  
33 55.6 8.3 392 4 US-09-325-932A-113 Sequence 113, App  
34 55.4 8.3 1553 3 US-08-821-994-60 Sequence 60, App  
35 54.8 8.2 1390 3 US-08-821-994-61 Sequence 61, App  
36 54.8 8.2 1441 3 US-08-821-994-63 Sequence 63, App  
37 54 8.1 1661 3 US-08-821-994-86 Sequence 86, App  
38 51.6 7.7 1102 3 US-08-821-994-86 Sequence 86, App  
39 50.6 7.6 1474 4 US-09-325-932A-101 Sequence 101, App  
40 50 7.5 743 3 US-08-821-994-37 Sequence 37, App  
41 50 7.5 1434 3 US-08-821-994-62 Sequence 62, App  
42 50 7.5 1474 3 US-08-821-994-64 Sequence 64, App  
43 49.2 7.4 1507 3 US-08-821-994-75 Sequence 75, App  
44 48.6 7.3 1577 3 US-08-821-994-59 Sequence 59, App  
45 48.4 7.3 468 3 US-08-821-994-36 Sequence 36, App

#### ALIGNMENTS

RESULT 1  
US-07-945-288-1  
; Sequence 1, Application US/07945288  
; Patent No. 5433948  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Wayne R.  
; APPLICANT: Chua, Kaw-Yan  
; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/945,288  
; FILING DATE: 19920910  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 580,655  
; FILING DATE: 11 SEPTEMBER 1990  
; APPLICATION NUMBER: 458,642  
; FILING DATE: 13 FEBRUARY 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MANDRACURAS, AMY E.  
; REGISTRATION NUMBER: P36,207  
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 834 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..738  
US-07-945-288-1

Query Match 99.8%; Score 664.4; DB 1; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 4.2e-209;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 ACTAACGCTGCAGTATCAATGGAATGCTCCAGTCAATGATTCGACAAATGCGA 60
Db 70 ACTAACGCTGCAGTATCAATGGAATGCTCCAGTCAATGATTCGACAAATGCGA 129
Qy 61 ACTGTCACTCCCAATTCCTATGCAAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 120
Db 130 ACTGTCACTCCCAATTCCTATGCAAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 189
Qy 121 CGCGCAACTGAATCAGCTTATTTGGCTCACCAGGTAATCAATCATGATTCGTGAACAA 180
Db 190 CGCGCAACTGAATCAGCTTATTTGGCTCACCAGGTAATCAATCATGATTCGTGAACAA 249
Qy 181 GAATTAGTTCGATTCGCTTCCCAACACGGTTCATGCTGATACCATTCACGTGTATT 240
Db 250 GAATTAGTTCGATTCGCTTCCCAACACGGTTCATGCTGATACCATTCACGTGTATT 309
Qy 241 GAATPACATCCAAATATGTTGGTCCGTAACGTAATCAATCATGATTCGTGAACAA 300
Db 310 GAATPACATCCAAATATGTTGGTCCGTAACGTAATCAATCATGATTCGTGAACAA 369
Qy 301 CAATCATGCGGACCAACCAATGCTGCTCCAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 360
Db 370 CAATCATGCGGACCAACCAATGCTGCTCCAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 429
Qy 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTGCGGTC 420
Db 430 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTGCGGTC 489
Qy 421 ATTATTGGATCAAGATTTAGACGATTCCTGATATGATGCGGCAATCATTCOA 480
Db 490 ATTATTGGATCAAGATTTAGACGATTCCTGATATGATGCGGCAATCATTCOA 549
Qy 481 CGGATATGTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 540
Db 550 CGGATATGTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 609
Qy 541 CAGGTGTCGATTTAGTTCGTAACCAAGTTCGATACCAATTCGATGCGGTAATGTT 600
Db 610 CAGGTGTCGATTTAGTTCGTAACCAAGTTCGATACCAATTCGATGCGGTAATGTT 669
Qy 601 TAGGTTATTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 660
Db 670 TAGGTTATTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 729
Qy 661 ATTCTC 666
Db 730 ATTCTC 735

```

## RESULT 2

US-08-462-831-1  
 ; Sequence 1, Application US/08462831  
 ; Patent No. 5552142  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
 ; DERMATOPHAGOIDES  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM: disk  
 ; MEDIUM TYPE: Floppy  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462.831
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..738
; US-08-462-831-1

```

Query Match 99.8%; Score 664.4; DB 1; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 4.2e-209;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 ACTAACGCTGCAGTATCAATGGAATGCTCCAGTCAATGATTCGACAAATGCGA 60
Db 70 ACTAACGCTGCAGTATCAATGGAATGCTCCAGTCAATGATTCGACAAATGCGA 129
Qy 61 ACTGTCACTCCCAATTCCTATGCAAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 120
Db 130 ACTGTCACTCCCAATTCCTATGCAAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 189
Qy 121 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATGATTCGTGAACAA 180
Db 190 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATGATTCGTGAACAA 249
Qy 181 GAATTAGTTCGATTCGCTTCCCAACACGGTTCATGCTGATACCATTCACGTGTATT 240
Db 250 GAATTAGTTCGATTCGCTTCCCAACACGGTTCATGCTGATACCATTCACGTGTATT 309
Qy 241 GAATACATCCAAATATGTTGGTCCCAAGAGGCTTACTATCGATACGTTGCACGAGAA 300
Db 310 GAATACATCCAAATATGTTGGTCCCAAGAGGCTTACTATCGATACGTTGCACGAGAA 369
Qy 301 CAATCATGCGGACCAACCAATGCTGCTCCAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 360
Db 370 CAATCATGCGGACCAACCAATGCTGCTCCAGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT 429
Qy 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTGCGGTC 420
Db 430 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTGCGGTC 489
Qy 421 ATTATTGGATCAAGATTTAGACGATTCCTGATATGATGCGGCAATCATTCOA 480
Db 490 ATTATTGGATCAAGATTTAGACGATTCCTGATATGATGCGGCAATCATTCOA 549
Qy 481 CGGATATGTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 540
Db 550 CGGATATGTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 609
Qy 541 CAGGTGTCGATTTAGTTCGTAACCAAGTTCGATACCAATTCGATGCGGTAATGTT 600
Db 610 CAGGTGTCGATTTAGTTCGTAACCAAGTTCGATACCAATTCGATGCGGTAATGTT 669
Qy 601 TAGGTTATTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 660
Db 670 TAGGTTATTTGTCGCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTACGCA 729
Qy 661 ATTCTC 666
Db 730 ATTCTC 735

```





FILING DATE: 13 FEBRUARY 1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: MANDRAGOURAS, AMY E.  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: IPC-0100C (IMI-024)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 834 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..738  
 US-08-461-441-1

Query Match 99.8%; Score 664.4; DB 1; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 4.2e-209;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTGGCAGCAATGCCA	60
Db	70	ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTGGCAGCAATGCCA	129
Qy	61	ACTGTCACTCCCATTCGTATGCAAGGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT	120
Db	130	ACTGTCACTCCCATTCGTATGCAAGGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT	189
Qy	121	GCGCCCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATGATCTTGCCTGAACAA	180
Db	190	GCGCCCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATGATCTTGCCTGAACAA	249
Qy	181	GAATTAGTCGATTGTGCTTCCCAACACGGTGTGTCATGGTGATACCAATTCACCGTGTATT	240
Db	730	ATTCTC 666	
Qy	730	ATTCTC 735	

RESULT 5  
 US-08-482-142-1  
 Sequence 1, Application US/08482142  
 Patent No. 5820862  
 GENERAL INFORMATION:  
 APPLICANT: Garman, Richard  
 APPLICANT: Greenstein, Julia  
 APPLICANT: Kuo, Mei-chang  
 APPLICANT: Rogers, Bruce  
 APPLICANT: Franzen, Henry  
 APPLICANT: Chen, Xian  
 APPLICANT: Evans, Sean  
 APPLICANT: Shaked, Ze'ev  
 TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 NUMBER OF SEQUENCES: 207  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 STREET: 610 LINCOLN STREET  
 CITY: WALTHAM  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02154  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII TEXT  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/482,142  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/445,307  
 FILING DATE: 07 June 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CRAIG, ANNE I.  
 REGISTRATION NUMBER: 32,976  
 REFERENCE/DOCKET NUMBER: 017.6US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 466-6000  
 TELEFAX: (617) 466-6040  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 834 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..738  
 US-08-482-142-1

Query Match 99.8%; Score 664.4; DB 1; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 4.2e-209;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTGGCAGCAATGCCA	60
Db	70	ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTGGCAGCAATGCCA	129
Qy	61	ACTGTCACTCCCATTCGTATGCAAGGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT	120
Db	130	ACTGTCACTCCCATTCGTATGCAAGGAGGCTGGTTCATGTTGGGCTTCTCTGGTGT	189
Qy	121	GCGCCCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATGATCTTGCCTGAACAA	180
Db	190	GCGCCCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATGATCTTGCCTGAACAA	249
Qy	181	GAATTAGTCGATTGTGCTTCCCAACACGGTGTGTCATGGTGATACCAATTCACCGTGTATT	240
Db	250	GAATTAGTCGATTGTGCTTCCCAACACGGTGTGTCATGGTGATACCAATTCACCGTGTATT	309

241 GAATACATCAACATAAATGGTGTCTGCTCAAGAAAGCTACTATCGATAGCTTGCACGAGAA 300  
Db  
310 GAATACATCAACATAAATGGTGTCTGCTCAAGAAAGCTACTATCGATAGCTTGCACGAGAA 369  
Qy  
301 CAATCATGCGGACGACCAAAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db  
370 CAATCATGCGGACGACCAAAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 429  
Qy  
361 CCACCAAAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
Db  
430 CCACCAAAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 489  
Qy  
421 ATTATGGCATCAAGATTTAGACGATTCCTGATTTATGATGGCGCAACAATCAATCAAA 480  
Db  
490 ATTATGGCATCAAGATTTAGACGATTCCTGATTTATGATGGCGCAACAATCAATCAAA 549  
Qy  
481 CCGGATAATGGTTACCAACCAACTATCAGCGTGTCAACATTTGGTGTACAGTAACGCA 540  
Db  
550 CCGGATAATGGTTACCAACCAACTATCAGCGTGTCAACATTTGGTGTACAGTAACGCA 609  
Qy  
541 CAAGGTGTGATTTATGGATTCGTAGCAACAGTTGGGATACCAATTTGGGTTGATAATGGT 600  
Db  
610 CAAGGTGTGATTTATGGATTCGTAGCAACAGTTGGGATACCAATTTGGGTTGATAATGGT 669  
Qy  
601 TAGCGTTATTTGCTGCCAACAATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db  
670 TAGCGTTATTTGCTGCCAACAATGATGATGATGATGATGATGATGATGATGATGATGAT 729  
Qy  
661 ATTCTC 666  
Db  
730 ATTCTC 735

## RESULT 6

US-08-478-572-1  
; Sequence 1, Application US/08478572  
; Patent No. 5968526  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/478,572  
; FILING DATE: 07-June-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/445,307  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: 017,605

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 834 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..738  
; US-08-478-572-1

Query Match 99.8%; Score 664.4; DB 2; Length 834;

Best Local Similarity 99.8%; Pred. No. 4.2e-209;

Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGACAAATCGCA 60  
Db 70 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGACAAATCGCA 129  
Qy 61 ACTGTCACTCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTTCTCTGTGTT 120  
Db 130 ACTGTCACTCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTTCTCTGTGTT 189  
Qy 121 GCCCAACTGAATCAGCTTATTTGGCTCACCCTGTAATCAATCATTTGGATCTTGTGAACAA 180  
Db 190 GCCCAACTGAATCAGCTTATTTGGCTCACCCTGTAATCAATCATTTGGATCTTGTGAACAA 249  
Qy 181 GAATTAGTCGATTGCTTCCCAACACGGTTGTCATGGTGCATACCATTCACCGTGGTATT 240  
Db 250 GAATTAGTCGATTGCTTCCCAACACGGTTGTCATGGTGCATACCATTCACCGTGGTATT 309  
Qy 241 GAATACATCCAAACATAATGCTGTCCTCAAGAAAGTACTATCGATACGTTGCACGAGAA 300  
Db 310 GAATACATCCAAACATAATGCTGTCCTCAAGAAAGTACTATCGATACGTTGCACGAGAA 369  
Qy 301 CAATCATGCGGACGACCAAAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 370 CAATCATGCGGACGACCAAAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 429  
Qy 361 CCACCAAAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
Db 430 CCACCAAAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 489  
Qy 421 ATTATGGCATCAAGATTTAGACGATTCCTGATTTATGATGGCGCAACAATCAATCAAA 480  
Db 490 ATTATGGCATCAAGATTTAGACGATTCCTGATTTATGATGGCGCAACAATCAATCAAA 549  
Qy 481 CCGGATAATGGTTACCAACCAACTATCAGCGTGTCAACATTTGGTGTACAGTAACGCA 540  
Db 550 CCGGATAATGGTTACCAACCAACTATCAGCGTGTCAACATTTGGTGTACAGTAACGCA 609  
Qy 541 CAAGGTGTGATTTATGGATTCGTAGCAACAGTTGGGATACCAATTTGGGTTGATAATGGT 600  
Db 610 CAAGGTGTGATTTATGGATTCGTAGCAACAGTTGGGATACCAATTTGGGTTGATAATGGT 669  
Qy 601 TAGCGTTATTTGCTGCCAACAATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db 670 TAGCGTTATTTGCTGCCAACAATGATGATGATGATGATGATGATGATGATGATGATGAT 729  
Qy 661 ATTCTC 666  
Db 730 ATTCTC 735

## RESULT 7

US-08-484-296-1  
; Sequence 1, Application US/08484296  
; Patent No. 6268491  
; GENERAL INFORMATION:

APPLICANT: Garman, Richard  
 APPLICANT: Greenstein, Julia  
 APPLICANT: Kuo, Mei-chang  
 APPLICANT: Rogers, Bruce  
 APPLICANT: Franzen, Henry  
 APPLICANT: Chen, Xian  
 APPLICANT: Evans, Sean  
 APPLICANT: Shaked, Ze'ev  
 TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 NUMBER OF SEQUENCES: 207  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 STREET: 610 LINCOLN STREET  
 CITY: WALTHAM  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02154  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII TEXT  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/484,296  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/445,307  
 FILING DATE: 07 June 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CRAIG, ANNE I.  
 REGISTRATION NUMBER: 32,976  
 REFERENCE/DOCKET NUMBER: 017,6US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 466-6000  
 TELEFAX: (617) 466-6040  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 834 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..738  
 PS-08-484-296-1

Qy	301	CAATCATGCGCAACCAAAATGCAACAAGTTTCGGTATCTCAAACTATTGCCAAATTTTAC	360
Db	370	CAATCATGCGCAACCAAAATGCAACAAGTTTCGGTATCTCAAACTATTGCCAAATTTTAC	429
Qy	361	CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACACAGCGCTATTGCCGTC	420
Db	430	CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACACAGCGCTATTGCCGTC	489
Qy	421	ATTATTGGCATCAAAAGATTTAGACGCATTCGGTCATTATGATGGCCGAAACAATCATTTCAA	480
Db	490	ATTATTGGCATCAAAAGATTTAGACGCATTCGGTCATTATGATGGCCGAAACAATCATTTCAA	549
Qy	481	CCGGAATATGTTACCAACCAAACTATCAGCTGTCAAGATTGTTGGTTACAGTAACGCA	540
Db	550	CCGGAATATGTTACCAACCAAACTATCAGCTGTCAAGATTGTTGGTTACAGTAACGCA	609
Qy	541	CAAGGTGTCGATTATTGGATGTCAGAAACAGATTGGGATACCAATTTGGGTGATATATGTT	600
Db	610	CAAGGTGTCGATTATTGGATGTCAGAAACAGATTGGGATACCAATTTGGGTGATATATGTT	669
Qy	601	TACGGTTATTTTGTGCGCAACATCGATTTGATGATTAAGAAATATCATATGTTGTC	660
Db	670	TACGGTTATTTTGTGCGCAACATCGATTTGATGATTAAGAAATATCATATGTTGTC	729
Qy	661	ATTCTC 666	
Db	730	ATTCTC 735	
RESULT 8			
PCT-US93-08518-1			
; Sequence 1, Application PC/TUS9308518			
; GENERAL INFORMATION:			
; APPLICANT:			
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM			
; TITLE OF INVENTION: DERMATOPHAGOIDES			
; NUMBER OF SEQUENCES: 13			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: LAHIVE & COCKFIELD			
; STREET: 60 STATE STREET, SUITE 510			
; CITY: BOSTON			
; STATE: MA			
; COUNTRY: USA			
; ZIP: 02109			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: ASCII TEXT			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: PCT/US93/08518			
; FILING DATE:			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/945,288			
; FILING DATE: 10 SEPTEMBER 1992			
; ATTORNEY/AGENT INFORMATION:			
; NAME: MANDRAGOURAS, AMY E.			
; REGISTRATION NUMBER: 36,207			
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (617) 227-7400			
; TELEFAX: (617) 227-5941			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 834 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: CDNA			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 1..738			

PCT-US93-08518-1

Query Match 99.8%; Score 664.4; DB 5; Length 834;  
Best Local Similarity 99.8%; Pred. No. 4.2e-209;  
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAAGCGCTGAGTATCAATGAATGCTCCAGCTGAAATCGATTTCGCAATGCGA 60  
DB 70 ACTAAGCGCTGAGTATCAATGAATGCTCCAGCTGAAATCGATTTCGCAATGCGA 129  
QY 61 ACTGTCACATCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 120  
DB 130 ACTGTCACATCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 189  
QY 121 GCGGCACTGAATCAGCTTATTTGGCTCACGTAATCAATCAATTCGATCTTGCTGAACAA 180  
DB 190 GCGGCACTGAATCAGCTTATTTGGCTCACGTAATCAATCAATTCGATCTTGCTGAACAA 249  
QY 181 GAATTAGTCGATTGTGCTTCCCAACACAGGTTGTGATGATACCAATCCACGTCGTATT 240  
DB 250 GAATTAGTCGATTGTGCTTCCCAACACAGGTTGTGATGATACCAATCCACGTCGTATT 309  
QY 241 GAATACATCCACATATTCGTGCTGCTCAAGAAAGCTACTATCGATAGCTTGACAGAA 300  
DB 310 GAATACATCCACATATTCGTGCTGCTCAAGAAAGCTACTATCGATAGCTTGACAGAA 369  
QY 301 CAATCATGCGACCAACCAATGCAAACTGTTTGGTATCTCAAACTATTTGCCAAATTTAC 360  
DB 370 CAATCATGCGACCAACCAATGCAAACTGTTTGGTATCTCAAACTATTTGCCAAATTTAC 429  
QY 361 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGTCAAAACCAACGAGCTATTCGCGTC 420  
DB 430 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGTCAAAACCAACGAGCTATTCGCGTC 489  
QY 421 ATTATGGCATCAAGATTTAGACGATTCGCTCATTTATGATGCGCGACATATTCATCA 480  
DB 490 ATTATGGCATCAAGATTTAGACGATTCGCTCATTTATGATGCGCGACATATTCATCA 549  
QY 481 CCGGATAATGGTTACCAACCAACTATACGCTGTCAACATTTGTTGGTTACAGTAACGCA 540  
DB 550 CCGGATAATGGTTACCAACCAACTATACGCTGTCAACATTTGTTGGTTACAGTAACGCA 609  
QY 541 CAAAGTGTGATATTTGGATCGTACGAACAGTTGGGATACCAATGGGATGATATGTT 600  
DB 610 CAAAGTGTGATATTTGGATCGTACGAACAGTTGGGATACCAATGGGATGATATGTT 669  
QY 601 TACGGTTATTTGCTGCCAACATCGATTTGATGATGATGATGATGATGATGATGATGAT 660  
DB 670 TACGGTTATTTGCTGCCAACATCGATTTGATGATGATGATGATGATGATGATGATGAT 729  
QY 661 ATTCTC 666  
DB 730 ATTCTC 735

## RESULT 9

US-07-945-288-9

Sequence 9, Application US/07945288

Patent No. 5433948

GENERAL INFORMATION:

APPLICANT: Thomas, Wayne R.

APPLICANT: Chua, Kow-Yan

TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM

DERMATOPHAGOIDES (HOUSE DUST MITES)

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE &amp; COCKFIELD

STREET: 60 STATE STREET, SUITE 510

CITY: BOSTON

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/945,288  
FILING DATE: 19920910  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: P36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1172 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..738  
US-07-945-288-9

Query Match 99.8%; Score 664.4; DB 1; Length 1172;  
Best Local Similarity 99.8%; Pred. No. 5.1e-209;  
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAAGCGCTGAGTATCAATGAATGCTCCAGCTGAAATCGATTTCGCAATGCGA 60  
DB 344 ACTAAGCGCTGAGTATCAATGAATGCTCCAGCTGAAATCGATTTCGCAATGCGA 403  
QY 61 ACTGTCACATCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 120  
DB 404 ACTGTCACATCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 463  
QY 121 GCGGCACTGAATCAGCTTATTTGGTCAACGTAATCAATCAATTCGATGTAACAA 180  
DB 464 GCGGCACTGAATCAGCTTATTTGGTCAACGTAATCAATCAATTCGATGTAACAA 523  
QY 181 GAATTAGTCGATTGTGCTTCCCAACACAGGTTGTGATGATGATGATGATGATGATGAT 240  
DB 524 GAATTAGTCGATTGTGCTTCCCAACACAGGTTGTGATGATGATGATGATGATGATGAT 583  
QY 241 GAATACATCCCAACATATTCGTCTCAAGAAAGCTACTATCGATACGTTGCACGAGAA 300  
DB 584 GAATACATCCCAACATATTCGTCTCAAGAAAGCTACTATCGATACGTTGCACGAGAA 643  
QY 301 CAATCATGCGACCAACCAATGCAAACTGTTGGTATCTCAAACTATTCGCAATTTAC 360  
DB 644 CAATCATGCGACCAACCAATGCAAACTGTTGGTATCTCAAACTATTCGCAATTTAC 703  
QY 361 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCAACGAGCTATTCGCGTC 420  
DB 704 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCAACGAGCTATTCGCGTC 763  
QY 421 ATTATGGCATCAAGATTTAGACGATTCGCTCATTTATGATGCGCGACAAATTCATCAA 480  
DB 764 ATTATGGCATCAAGATTTAGACGATTCGCTCATTTATGATGCGCGACAAATTCATCAA 823  
QY 481 GCGGATATGTTTACCAACCAACTATCAGCTGTCAACATTTGTTGGTTACAGTAACGCA 540  
DB 824 GCGGATATGTTTACCAACCAACTATCAGCTGTCAACATTTGTTGGTTACAGTAACGCA 883  
QY 541 CAAAGTGTGATATTTGGATCGTACGAACAGTTGGGATACCAATTCGCGGTGATAATGTT 600

Db 884 CAAGGTGCGATTATGCGATCGTACGAAACAGTTGGGATACCAATTGGGTGATAATGGT 943  
Qy 601 TACGGTTATTTGCTGCGCAACATCGATTTCATGATGATTGAAGAAATCCATATGTTGTC 660  
Db 944 TACGGTTATTTGCTGCGCAACATCGATTTCATGATGATTGAAGAAATCCATATGTTGTC 1003  
Qy 661 ATTCTC 666  
Db 1004 ATTCTC 1009

RESULT 10  
US-08-462-831-9  
; Sequence 9, Application US/08462831  
; Patent No. 5552142  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,831  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION NUMBER:  
; FILING DATE: 10 SEPTEMBER 1992  
; APPLICATION NUMBER: US 07/945,288  
; FILING DATE: 11 SEPTEMBER 1990  
; APPLICATION NUMBER: US 580,655  
; FILING DATE: 13 FEBRUARY 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MANDRAGOURAS, AMY E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1172 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..738  
; US-08-462-831-9

Query Match 99.8%; Score 664.4; DB 1; Length 1172;  
Best Local Similarity 99.8%; Pred. No. 5.1e-209;  
Matches 665; Conservative 1; Indels 0; Gaps 0;

Qy 1 ACTAACCCCTCGAGTATCAATGGAATGCTCCAGCTCAAAATCCATTGCGACAAATGCCA 60  
Db 344 ACTAACCCCTCGAGTATCAATGGAATGCTCCAGCTCAAAATCCATTGCGACAAATGCCA 403  
Qy 61 ACTGTACTCCCATTCGATGATGCAAGGAGGCTGGTTCATGTTGGCTTTCTCTGTTGTT 120

Db 404 ACTGTACTCCCATTCGATGATGCAAGGAGGCTGGTTCATGTTGGCTTTCTCTGTTGTT 463  
Qy 121 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATCAATGATCTTGTGAAACAA 180  
Db 464 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATGATCTTGTGAAACAA 523  
Qy 181 GAATTAGTCGATTGTTGCTTCCCAACACGGTTGTCATGGTGATACCATTCACGTTGGTATT 240  
Db 524 GAATTAGTCGATTGTTGCTTCCCAACACGGTTGTCATGGTGATACCATTCACGTTGGTATT 583  
Qy 241 GAATACATCCCAACATAATGTTGCTCCCAAGAAAGTACTATCGATACGTTGCAACGAGAA 300  
Db 584 GAATACATCCCAACATAATGTTGCTCCCAAGAAAGTACTATCGATACGTTGCAACGAGAA 643  
Qy 301 CAATCATGCCGACCAACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCAAAATTTAC 360  
Db 644 CAATCATGCCGACCAACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCAAAATTTAC 703  
Qy 361 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACACGCGTATTGCCGTC 420  
Db 704 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACACGCGTATTGCCGTC 763  
Qy 421 ATTATTGGCATCAAAGATTATTAGACGCAATTCGGTCATTATGATGCGCCGAAACAATTCAA 480  
Db 764 ATTATTGGCATCAAAGATTATTAGACGCAATTCGGTCATTATGATGCGCCGAAACAATTCAA 823  
Qy 481 CGCATATGTTTACCAACCAAACTATCAGCTGTCAACATTTGTTGTTTACGATTAACGCA 540  
Db 824 CGCATATGTTTACCAACCAAACTATCAGCTGTCAACATTTGTTGTTTACGATTAACGCA 883  
Qy 541 CAAGTGTCGATTATTGGATCGTACGAAACAGTTGGGATACCAATTCGGGTGATAATGCT 600  
Db 884 CAAGTGTCGATTATTGGATCGTACGAAACAGTTGGGATACCAATTCGGGTGATAATGCT 943  
Qy 601 TACGGTTATTTGCTGCGCAACATCGATTTCATGATGATTTGAAGAAATTCATATGTTGTC 660  
Db 944 TACGGTTATTTGCTGCGCAACATCGATTTCATGATGATTTGAAGAAATTCATATGTTGTC 1003  
Qy 661 ATTCTC 666  
Db 1004 ATTCTC 1009

RESULT 11  
US-08-461-809-9  
; Sequence 9, Application US/08461809  
; Patent No. 5770202  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,809  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/945,288  
; FILING DATE: 10 SEPTEMBER 1992  
; APPLICATION NUMBER: US 580,655  
; FILING DATE: 11 SEPTEMBER 1990

```

APPLICATION NUMBER: US 458,642
FILING DATE: 13 FEBRUARY 1990
ATTORNEY/AGENT INFORMATION:
NAME: MANDRAGOURAS, AMY E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1..738
US-08-461-809-9

Query Match
Best Local Similarity 99.8%; Score 664.4; DB 1; Length 1172;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAACGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60
Db 344 ACTAACGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 403
QY 61 ACTGTCACCTCCATTCGATGAAGAGGCTGTGTTCAATGTTGGCTTCTCTGGTGT 120
Db 404 ACTGTCACCTCCATTCGATGAAGAGGCTGTGTTCAATGTTGGCTTCTCTGGTGT 463
QY 404 ACTGTCACCTCCATTCGATGAAGAGGCTGTGTTCAATGTTGGCTTCTCTGGTGT 463
QY 121 GCGCAACTGAATCAGCTTATTTGGCTCCACCGTGAATCGAATTCGACAAATGCGA 180
Db 464 GCGCAACTGAATCAGCTTATTTGGCTCCACCGTGAATCGAATTCGACAAATGCGA 523
QY 181 GAATTAGTCGATTCGCTTCCCAACACGGTGTGTCATGTCGATACCATTCACCGTGTAT 240
Db 524 GAATTAGTCGATTCGCTTCCCAACACGGTGTGTCATGTCGATACCATTCACCGTGTAT 583
QY 241 GAATACATCCACATATATGTTGTCGTCACCAAGAGCTTACTATCGATACGTTGCCAGAA 300
Db 584 GAATACATCCACATATATGTTGTCGTCACCAAGAGCTTACTATCGATACGTTGCCAGAA 643
QY 301 CAATCATGCGACGACCAAAATGCACAAACGTTTCGTTATCTCAAACTATTGCCAAATTTAC 360
Db 644 CAATCATGCGACGACCAAAATGCACAAACGTTTCGTTATCTCAAACTATTGCCAAATTTAC 703
QY 361 CCACCAATGCAACAAATTCGTAAGCTTTGGCTCAACCCACAGCGCTATTGCCGTC 420
Db 704 CCACCAATGCAACAAATTCGTAAGCTTTGGCTCAACCCACAGCGCTATTGCCGTC 763
QY 421 ATTATGGCATCAAGATTAGACGCTATTCGCTCATTTATGATGCGCGCAACATCATTCAA 480
Db 764 ATTATGGCATCAAGATTAGACGCTATTCGCTCATTTATGATGCGCGCAACATCATTCAA 823
QY 481 CCGGATAATGTTACCAACCAAACTATACCGTGTCAACATTTGTTGTTACAGTAACGCA 540
Db 824 CCGGATAATGTTTACCAACCAAACTATACCGTGTCAACATTTGTTGTTACAGTAACGCA 883
QY 541 CAAGGTGTCGATTATTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTATATGTT 600
Db 884 CAAGGTGTCGATTATTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTATATGTT 943
QY 601 TAGGGTTATTTGCTGCCAACTCGATTGATGATGATTGAAGAATATCCATATGTTGTC 660
Db 944 TAGGGTTATTTGCTGCCAACTCGATTGATGATGATTGAAGAATATCCATATGTTGTC 1003
QY 661 ATTCTC 666
Db 1004 ATTCTC 1009

```

```

RESULT 12
US-08-461-441-9
; Sequence 9, Application US/08461441
; Patent No. 5773002
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,441
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..738
; US-08-461-441-9

Query Match
Best Local Similarity 99.8%; Score 664.4; DB 1; Length 1172;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAACGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60
Db 344 ACTAACGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 403
QY 61 ACTGTCACCTCCATTCGATGAAGAGGCTGTGTTCAATGTTGGCTTCTCTGGTGT 120
Db 404 ACTGTCACCTCCATTCGATGAAGAGGCTGTGTTCAATGTTGGCTTCTCTGGTGT 463
QY 121 GCGCAACTGAATCAGCTTATTTGGCTCCACCGTGAATCGAATTCGACAAATGCGA 180
Db 464 GCGCAACTGAATCAGCTTATTTGGCTCCACCGTGAATCGAATTCGACAAATGCGA 523
QY 181 GAATTAGTCGATTCGCTTCCCAACACGGTGTGTCATGTCGATACCATTCACCGTGTAT 240
Db 524 GAATTAGTCGATTCGCTTCCCAACACGGTGTGTCATGTCGATACCATTCACCGTGTAT 583
QY 241 GAATACATCCACATATATGTTGTCGTCACCAAGAGCTTACTATCGATACGTTGCCAGAA 300

```

Db 584 GAATACATCAACATAATGCTGCTGCTCAAGAAAGCTACTATCGATACGTTGCACGAGAA 643  
QY 301 CAATCATGCGCGACCAACAAATGCAACACGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 644 CAATCATGCGCGACCAACAAATGCAACACGTTTCGGTATCTCAAACTATTGCCAAATTTAC 703  
QY 361 CCACCAATGCAACAAATGCTGAAGCTTTGGCTCAAAACCCACACGCGTATTGCCGTC 420  
Db 704 CCACCAATGCAACAAATGCTGAAGCTTTGGCTCAAAACCCACACGCGTATTGCCGTC 763  
QY 421 ATTATGCGCATCAAGATTTAGACGCTATTCGGTCAATTCATGATGCGCGAAACATTCATCA 480  
Db 764 ATTATGCGCATCAAGATTTAGACGCTATTCGGTCAATTCATGATGCGCGAAACATTCATCA 823  
QY 481 CGCGATATGTTTACCAACCAACTATCACGCTGTCAACATTTGGTTCAGTAACGCA 540  
Db 824 CGCGATATGTTTACCAACCAACTATCACGCTGTCAACATTTGGTTCAGTAACGCA 883  
QY 541 CAGGTGTCGATTTATGCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGTT 600  
Db 884 CAGGTGTCGATTTATGCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGTT 943  
QY 601 TAGCGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGATATCCATATGTTGTC 660  
Db 944 TAGCGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGATATCCATATGTTGTC 1003  
QY 661 ATTCTC 666  
Db 1004 ATTCTC 1009

## RESULT 13

PCT-US93-08518-9

; Sequence 9, Application PC/TUS9308518

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM

; TITLE OF INVENTION: DERMATOPHAGOIDES

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE &amp; COCKFIELD

; STREET: 60 STATE STREET, SUITE 510

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/08518

; FILING DATE:

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 07/945,288

; FILING DATE: 10 SEPTEMBER 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: MANDRAGOURAS, AMY E.

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1172 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..738  
PCT-US93-08518-9

Query Match 99.8%; Score 664.4; DB 5; Length 1172;

Best Local Similarity 99.8%; Pred. No. 5.1e-209; Indels 0; Gaps 0;  
Matches 665; Conservative 0; Mismatches 1;

QY 1 ACTAACGCTGCACTATCAATGGAAATGCTCCAGCTCAAAATCGAATTCGACAAATGCGA 60  
Db 344 ACTAACGCTGCACTATCAATGGAAATGCTCCAGCTCAAAATCGAATTCGACAAATGCGA 403  
QY 61 ACTGTCATCCCATTCGATGTCAGAGAGGCTGGTTCATGTTGGGCTTCTCTGCTGTT 120  
Db 404 ACTGTCATCCCATTCGATGTCAGAGAGGCTGGTTCATGTTGGGCTTCTCTGCTGTT 463  
QY 121 GCCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGTGAACA 180  
Db 464 GCCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGTGAACA 523  
QY 181 GAATAGTCGATTTGCTTCCCAACACGTTGTGTCATGTCATGATACCATTCACGCTGATT 240  
Db 524 GAATAGTCGATTTGCTTCCCAACACGTTGTGTCATGTCATGATACCATTCACGCTGATT 583  
QY 241 GAATACATCCCAACATAATGCTGTCCTCAAGAAAGCTACTATCGATACGTTGCACGAGAA 300  
Db 584 GAATACATCCCAACATAATGCTGTCCTCAAGAAAGCTACTATCGATACGTTGCACGAGAA 643  
QY 301 CAATCATGCGGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTCGCAAAATTTAC 360  
Db 644 CAATCATGCGGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTCGCAAAATTTAC 703  
QY 361 CCACCAATGCAACAAATTTGCTGGAAGCTTTGGCTCAAAACCCACACGCGTATTGCCGTC 420  
Db 704 CCACCAATGCAACAAATTTGCTGGAAGCTTTGGCTCAAAACCCACACGCGTATTGCCGTC 763  
QY 421 ATTATGCGATCAAGATTTAGAGCATTCGGTATTCATGATGCGCGAAACATCATTTCAA 480  
Db 764 ATTATGCGATCAAGATTTAGAGCATTCGGTATTCATGATGCGCGAAACATCATTTCAA 823  
QY 481 CGCGATATGTTTACCAACCAACTATCACGCTGTCAACATTTGTTGTTACAGTAACGCA 540  
Db 824 CGCGATATGTTTACCAACCAACTATCACGCTGTCAACATTTGTTGTTACAGTAACGCA 883  
QY 541 CAGGTGTCGATTTATGCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGTT 600  
Db 884 CAGGTGTCGATTTATGCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGTT 943  
QY 601 TAGCGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGATATCCATATGTTGTC 660  
Db 944 TAGCGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGATATCCATATGTTGTC 1003  
QY 661 ATTCTC 666  
Db 1004 ATTCTC 1009

## RESULT 14

US-08-460-040-1

; Sequence 1, Application US/08460040

; Patent No. 6071522

; GENERAL INFORMATION:

; APPLICANT: Thomas, Wayne R.

; TITLE OF INVENTION: Cloning of Mite Allergens

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE &amp; COCKFIELD

; STREET: 60 State Street, suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109-1875



COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/450,040  
 FILING DATE: 2-JUNE-95  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/162,722  
 FILING DATE: 8-NOV-93  
 APPLICATION NUMBER: 07/458,642  
 FILING DATE: 13-FEB-90  
 APPLICATION NUMBER: PCT/AUS8/00195  
 FILING DATE: 17-JUNE-88  
 APPLICATION NUMBER: PI 2523/87  
 FILING DATE: 18-JUNE-87  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amy E. Mandragouras  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: IMI-021CN2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 TELEFAX: (617)227-5941  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 857 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..736  
 FEATURE:  
 NAME/KEY: mat.peptide  
 LOCATION: 70..736  
 US-08-460-040-1

Query Match 95.4%; Score 635.2; DB 3; Length 857;  
 Best Local Similarity 99.0%; Pred. No. 1.8e-199;  
 Matches 661; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 1 ACTACGCTGAGTATCAATGGAATGCTCAGCTGAAATCGATTTCGCGCAAAATGCGA 60  
 DB 70 ACTACGCTGAGTATCAATGGAATGCTCAGCTGAAATCGATTTCGCGCAAAATGCGA 129  
 QY 61 ACTGTCACCTCCCAATTCGATGCAAGAGGCTGTGTTTCATGTTGGCTTTCTCTGGTGT 120  
 DB 130 ACTGTCACCTCCCAATTCGATGCAAGAGGCTGTGTTTCATGTTGGCTTTCTCTGGTGT 189  
 QY 121 GCCGCAACTGAATCAGCTTATTTGGCTCAGCTATCAATCAATGGATCTTGCTGAACAA 180  
 DB 190 GCCGCAACTGAATCAGCTTATTTGGCTCAGCTATCAATCAATGGATCTTGCTGAACAA 249  
 QY 181 GAATTAGTCGATGTTGCTTCCCAACACCGTTGCTATGTTGATACCAATTCACGTTGAT 240  
 DB 250 GAATTAGTCGATGTTGCTTCCCAACACCGTTGCTATGTTGATACCAATTCACGTTGAT 309  
 QY 241 GAATACATCCACATAATGTTGCTCCAGAAAGCTACTATCGATAGTTGCGACGAGAA 300  
 DB 310 GAATACATCCACATAATGTTGCTCCAGAAAGCTACTATCGATAGTTGCGACGAGAA 369  
 QY 301 CAATATGCGGACCAACCAATGCAACGTTTGGTATCTCAAACTATTGCAAAATTTAC 360  
 DB 370 CAATATGCGGACCAACCAATGCAACGTTTGGTATCTCAAACTATTGCAAAATTTAC 429  
 QY 361 CCACCAATGCAAAACAAATTCGTAAGCTTTGGCTCAAAACCAACAGCGCTATTGCGGCTC 420  
 DB 430 CCACCAATGCAAAACAAATTCGTAAGCTTTGGCTCAAAACCAACAGCGCTATTGCGGCTC 487  
 QY 421 ATTATGG--CATCAAGATTTAGACGCAATTCGCTCATTAATGATGCGCGCAACATCATTC 478

DB 488 ATTATTGGACCATCAAAAGATTTAGACGCAATTCGTCATTATGATGCGCAACATCATTC 547  
 QY 479 AACGCGATANTGTTTACCAACCAAACTATACGCTGTCACATTTGTTGTTACAGTAACG 538  
 DB 548 AACGCGATAATGTTTACCAACCAAACTATACGCTGTCACATTTGTTGTTACAGTAACG 607  
 QY 539 CACAAGGTGTCGATTATTGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATAATG 598  
 DB 608 CACAAGGTGTCGATTATTGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATAATG 667  
 QY 599 GTTACGCTATTTTGTGCGCAACATGATTTGATGATTTGAAGATATCCATATGTTG 658  
 DB 668 GTTACGCTATTTTGTGCGCAACATGATTTGATGATTTGAAGATATCCATATGTTG 727  
 QY 659 TCATTCTC 666  
 DB 728 TCATTCTC 735

RESULT 15  
 US-07-945-288-5  
 ; Sequence 5, Application US/07945288  
 ; Patent No. 5433948  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thomas, Wayne R.  
 ; APPLICANT: Chua, Kow-yan  
 ; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
 ; TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/945,288  
 ; FILING DATE: 19920910  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 580,655  
 ; FILING DATE: 11 SEPTEMBER 1990  
 ; APPLICATION NUMBER: 458,642  
 ; FILING DATE: 13 FEBRUARY 1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MANDRAGOURAS, AMY E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1072 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 36..1001  
 ; US-07-945-288-5

Query Match 70.5%; Score 469.8; DB 1; Length 1072;  
 Best Local Similarity 83.4%; Pred. No. 8e-145;  
 Matches 532; Conservative 0; Mismatches 106; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 20:26:03 ; Search time 3504.5 Seconds  
(without alignments)  
8987.005 Million cell updates/sec

Title: US-09-867-159A-1

Perfect score: 666

Sequence: 1 actaacgctgcagatataca.....atccatattgttcattctc 666

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hng.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	666	6	AX586972 Sequence
2	664.4	99.8	834	6	A73351 Sequence 1
3	664.4	99.8	834	6	AR013793 Sequence
4	664.4	99.8	834	6	AR047909 Sequence
5	664.4	99.8	834	6	AR080065 Sequence
6	664.4	99.8	834	6	I13095 Sequence 1
7	664.4	99.8	834	6	I25468 Sequence 1
8	664.4	99.8	857	6	BD001750 Tick allele
9	664.4	99.8	1172	6	AR013797 Sequence
10	664.4	99.8	1172	6	I13099 Sequence 9
11	664.4	99.8	1172	6	I25472 Sequence 9
12	661.2	99.3	666	6	AX459824 Sequence
13	661.2	99.3	669	6	AX114208 Sequence
14	661.2	99.3	669	6	AX114210 Sequence
15	661.2	99.3	849	6	CQ786568 Sequence
16	661.2	99.3	906	6	AX114205 Sequence
17	661.2	99.3	906	6	AX114207 Sequence
18	661.2	99.3	909	6	AX339354 Sequence
19	661.2	99.3	960	6	AX114202 Sequence

C	20	661.2	99.3	960	6	AX114204	Sequence
	21	661.2	99.3	1099	3	DP011695	U11695 Dermatophag
	22	661.2	99.3	1099	6	CQ797578	Sequence
	23	661.2	99.3	1099	6	CQ802486	Sequence
	24	659.6	99.0	1089	6	CQ797566	Sequence
	25	659.6	99.0	1089	6	CQ802474	Sequence
	26	658.2	98.8	951	6	AX022865	Sequence
	27	658.2	98.8	963	6	BD086203	Recombina
	28	658	98.8	963	6	AX022863	Sequence
	29	658	98.8	963	6	BD086202	Recombina
	30	656.4	98.6	963	6	AX022867	Sequence
	31	656.4	98.6	963	6	BD086204	Recombina
	32	652.2	97.9	669	6	AX114261	Sequence
C	33	652.2	97.9	669	6	AX114263	Sequence
	34	652.2	97.9	909	6	AX114256	Sequence
C	35	652.2	97.9	909	6	AX114258	Sequence
	36	649	97.4	666	6	AX114273	Sequence
C	37	649	97.4	666	6	AX114275	Sequence
	38	649	97.4	906	6	AX114268	Sequence
C	39	649	97.4	906	6	AX114270	Sequence
	40	644.2	96.7	906	6	AX114278	Sequence
C	41	644.2	96.7	906	6	AX114280	Sequence
	42	640.4	96.2	666	6	AX459752	Sequence
	43	640.4	96.2	666	6	AX459758	Sequence
	44	638.8	95.9	666	6	AX459754	Sequence
	45	637.2	95.7	666	6	AX459750	Sequence

#### ALIGNMENTS

RESULT 1  
AX586972  
LOCUS AX586972 666 bp DNA linear PAT 10-JAN-2003  
DEFINITION Sequence 1 from Patent WO02078736.  
ACCESSION AX586972  
VERSION AX586972.1 GI:27655851  
KEYWORDS  
SOURCE Dermatophagoides pteronyssinus (European house dust mite)  
ORGANISM Dermatophagoides pteronyssinus  
Bukaryota; Metazoa; Arthropoda; Chelicerata; Atachnida; Acari;  
Acariformes; Sarcopitiformes; Astigmata; Psoroptidia; Analgoidea;  
Pyroglyphidae; Dermatophagoides.  
REFERENCE 1  
AUTHORS Loria, E., Terrasse, G. and Trehin, Y.  
TITLE Antiallergic pharmaceutical composition  
JOURNAL Patent: WO 02078736-A 1 10-OCT-2002;  
Antialis (FR)  
FEATURES  
Location/Qualifiers  
source 1..666  
/organism="Dermatophagoides pteronyssinus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6956"

ORIGIN	Query Match	100.0%;	Score 666;	DB 6;	Length 666;
	Best Local Similarity	100.0%;	Pred. No. 1.3e-181;		
	Matches 666;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACTAACGCTGCAGTATCAATCGAAATGCTCCAGCTGAAATCGATTGCGACAAATGCCA	60		
Db	1	ACTAACGCTGCAGTATCAATCGAAATGCTCCAGCTGAAATCGATTGCGACAAATGCCA	60		
QY	61	ACTGTCACTCCCATTCGTTATCAAGAGGCTGCTGTTTCATGTTGGCTTCTCTGTTGT	120		
Db	61	ACTGTCACTCCCATTCGTTATCAAGAGGCTGCTGTTTCATGTTGGCTTCTCTGTTGT	120		
QY	121	GGCGCAACTGAATCAGCTTATTGGCTCACCCTTAATCAATCATTCGATCTTCTGTAACAA	180		
Db	121	GGCGCAACTGAATCAGCTTATTGGCTCACCCTTAATCAATCATTCGATCTTCTGTAACAA	180		
QY	181	GAATAGTCGATTGCTTCCCAACACGGTTGTCATGGTGATACCAATCCACGCTGTTATT	240		

```

Db      181 GAATTAGTCGATTGTGCTTCCCAACACGGTTGTGTCATGGTGATACCATTCACGTTGGTATT 240
QY      241 GAATACATCAACATAAATGGTGTGCTGTCGAAGAAAGCTACTATCGATACGTTTCACAGAA 300
Db      241 GAATACATCAACATAAATGGTGTGCTGTCGAAGAAAGCTACTATCGATACGTTTCACAGAA 300
QY      301 CAATCATGCGCGACCAACCAATGACAAACGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
Db      301 CAATCATGCGCGACCAACCAATGACAAACGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
QY      361 CCACAAATGCAACCAAAATTCGTGAAGCTTTGGTTCACAAACCCACAGCGCTATTGCCGTC 420
Db      361 CCACAAATGCAACCAAAATTCGTGAAGCTTTGGTTCACAAACCCACAGCGCTATTGCCGTC 420
QY      421 ATTATTGGCATCAAGATTTAGACGATTCGGTTCATTATGATGGCGCAACATTCATCAA 480
Db      421 ATTATTGGCATCAAGATTTAGACGATTCGGTTCATTATGATGGCGCAACATTCATCAA 480
QY      481 CGCGATAATGGTTTACCAACCAAACTATCAACGCTGTCAACATTTGGTTACAGTAACGCA 540
Db      481 CGCGATAATGGTTTACCAACCAAACTATCAACGCTGTCAACATTTGGTTACAGTAACGCA 540
QY      541 CAAGGTGTCAATTTATGCGTACCAACCAAACTATCAACGCTGTCAACATTTGGTTACAGTA 600
Db      541 CAAGGTGTCAATTTATGCGTACCAACCAAACTATCAACGCTGTCAACATTTGGTTACAGTA 600
QY      601 TACGGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGAATATCCATATGTTGTC 660
Db      601 TACGGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGAATATCCATATGTTGTC 660
QY      661 ATTCTC 666
Db      661 ATTCTC 666

RESULT 2
A73351 LOCUS 834 bp DNA linear PAT 15-OCT-1999
DEFINITION Sequence 1 from Patent WO942481.
ACCESSION A73351
VERSION A73351.1 GI:6064124
KEYWORDS unidentified
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 834)
AUTHORS Garman,R.D. and Kuo,M.
TITLE T CELL EPITOPES OF THE MAJOR ALLERGENS FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
JOURNAL IMMUNOLOGIC PHARMA CORP (US)
FEATURES
    source
        1..834
            /organism="unidentified"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32644"
            <1..738
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAB58491.1"
                /db_xref="GI:6064125"
                /translations="KRRFLMSAEAFELKLTQFDLNAETNACSINGNAPAEIDLQWRT
                VPIRMGGGCGSWAFSGVAESAYLAHRNQSLDLAEQLVDLSAQHCHGDTIPRG
                IEVIGHNVQVESYVYVAREQSCRPNACRQGISNYCOIYPNANKIREALAQTHSA
                IAVLIGKLDLAEFRHVDGRTLIQRDNGYQPNYHAVNIVGYSNAQGVYVIVENSMDTN
                WGDNGYGFPAANDLDMLEBYPYVIL"
CDS
ORIGIN
Query Match 99.8%; Score 664.4; DB 6; Length 834;
Best Local Similarity 99.8%; Pred. No. 3.9e-181;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTTCGACAAATGCGA 60

```

```

Db      70 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATTCGACAAATGCGA 129
QY      61 ACTGTCACCTCCCATTCGTATGCAAGAGGCTGTGGTTTCATGTTGGGCTTTCTCTGCTGTT 120
Db      130 ACTGTCACCTCCCATTCGTATGCAAGAGGCTGTGGTTTCATGTTGGGCTTTCTCTGCTGTT 189
QY      121 GCAGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCATTTGATCTTGTGAACAA 180
Db      190 GCAGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCATTTGATCTTGTGAACAA 249
QY      181 GAATTAGTCGATTGTGCTTCCCAACACGGTTGTGTCATGGTGATACCATTCACGTTGATT 240
Db      250 GAATTAGTCGATTGTGCTTCCCAACACGGTTGTGTCATGGTGATACCATTCACGTTGATT 309
QY      241 GAATACATCAACATTAATGTTGCTCCAAAGAAAGTACTATCGATACGTTGCACGAGAA 300
Db      310 GAATACATCAACATTAATGTTGCTCCAAAGAAAGTACTATCGATACGTTGCACGAGAA 369
QY      301 CAATCATGCGCGACCAACCAATGCAAACTGTCGATTCGTTTCCGATCTCAAACTATTGCCAA 360
Db      370 CAATCATGCGCGACCAACCAATGCAAACTGTCGATTCGTTTCCGATCTCAAACTATTGCCAA 429
QY      361 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Db      430 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 489
QY      421 ATTATTGGCATCAAGATTTAGACGATTCGGTTCATTATGATGGCGCAACATTCATTCAA 480
Db      490 ATTATTGGCATCAAGATTTAGACGATTCGGTTCATTATGATGGCGCAACATTCATTCAA 549
QY      481 CGCGATAATGGTTTACCAACCAAACTATCAACGCTGTCAACATTTGGTTACAGTAACGCA 540
Db      550 CGCGATAATGGTTTACCAACCAAACTATCAACGCTGTCAACATTTGGTTACAGTAACGCA 609
QY      541 CAAGGTGTCAATTTATGCGTACCAACCAAGTTCGGTACCAATTTGGGTTGATATGTT 600
Db      610 CAAGGTGTCAATTTATGCGTACCAACCAAGTTCGGTACCAATTTGGGTTGATATGTT 669
QY      601 TACGGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGAATATCCATATGTTGTC 660
Db      670 TACGGTTATTTGCTGCAACATCGATTTGATGATGATTGAAGAATATCCATATGTTGTC 729
QY      661 ATTCTC 666
Db      730 ATTCTC 735

RESULT 3
A73351 LOCUS 834 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 1 from patent US 5773002.
ACCESSION A73351
VERSION A73351.1 GI:3971247
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 834)
AUTHORS Thomas,W.R. and Chua,K.-Y.
TITLE Cloning and sequencing of allergens of dermatophagoides (house dust mite)
JOURNAL Patent: US 5773002-A 1 30-JUN-1998;
FEATURES
    Location/Qualifiers
        1..834
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
Query Match 99.8%; Score 664.4; DB 6; Length 834;
Best Local Similarity 99.8%; Pred. No. 3.9e-181;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATTCGACAAATGCGA 129

```

```

QY 1 ACTAAGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTCGCAAAATGCGA 60
|
|
|
Db 70 ACTAAGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTCGCAAAATGCGA 129
|
|
|
QY 61 ACTGTCACCTCCCATTCGATATGCAAGAGGCTGTGTTTCATGTTGGCTTCTCTGTTGTT 120
|
|
|
Db 130 ACTGTCACCTCCCATTCGATATGCAAGAGGCTGTGTTTCATGTTGGCTTCTCTGTTGTT 189
|
|
|
QY 121 GCCGCAATGAAATAGCTTATTTGGCTCACCGTAATCAATCAATCGATGCTTGAACAA 180
|
|
|
Db 190 GCCGCAATGAAATAGCTTATTTGGCTCACCGTAATCAATCAATCGATGCTTGAACAA 249
|
|
|
QY 181 GAATTAGTCGATTGCTTCCCAACACGGTTGCTATGCTGATACCATTCACCGTGGTATT 240
|
|
|
Db 250 GAATTAGTCGATTGCTTCCCAACACGGTTGCTATGCTGATACCATTCACCGTGGTATT 309
|
|
|
QY 241 GAATACATCCCAATATGTTGCTCCCAAGAAAGCTACTATCGATACGTTGCACGAGAA 300
|
|
|
Db 310 GAATACATCCCAATATGTTGCTCCCAAGAAAGCTACTATCGATACGTTGCACGAGAA 369
|
|
|
QY 301 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
|
|
|
Db 370 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 429
|
|
|
QY 361 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
|
|
|
Db 430 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 489
|
|
|
QY 421 ATTATGTCATCAAGATTTAGACGCTTCCGTCATTAATGATGCGCGCAACATCAATCAA 480
|
|
|
Db 490 ATTATGTCATCAAGATTTAGACGCTTCCGTCATTAATGATGCGCGCAACATCAATCAA 549
|
|
|
QY 481 CGCGATAATGTTTACCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTAACGCA 540
|
|
|
Db 550 CGCGATAATGTTTACCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTAACGCA 609
|
|
|
QY 541 CAAGGTGTCGATTATGATGCTACCAACAGTTGGGATACCAATTTGGGTTGATAATGTT 600
|
|
|
Db 610 CAAGGTGTCGATTATGATGCTACCAACAGTTGGGATACCAATTTGGGTTGATAATGTT 669
|
|
|
QY 601 TACGGTTATTTGCTGCCAACAATTCGATGATGATGATGATGATGATGATGATGATGATG 660
|
|
|
Db 670 TACGGTTATTTGCTGCCAACAATTCGATGATGATGATGATGATGATGATGATGATGATG 729
|
|
|
QY 661 ATTCTC 666
|
|
|
Db 730 ATTCTC 735

RESULT 4
AR047909
LOCUS AR047909 834 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5820862.
ACCESSION AR047909
VERSION AR047909.1 GI:5970252
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 834)
AUTHORS Garman,R.D., Greenstein,J.L., Kuo,M.-C., Rogers,B.L., Franzen,H.M.,
Chen,X., Evans,S. and Shaked,Z.
TITLE T cell epitopes of the major allergens from dermatophagoides (house
dust mite)
JOURNAL Patent: US 5820862-A 1 13-OCT-1998;
FEATURES Location/Qualifiers
source 1. 834
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 99.8%; Score 664.4; DB 6; Length 834;
Best Local Similarity 99.8%; Pred. No. 3.9e-181;

```

```

Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACTAAGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTCGCAAAATGCGA 60
|
|
|
Db 70 ACTAAGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTCGCAAAATGCGA 129
|
|
|
QY 61 ACTGTCACCTCCCATTCGATATGCAAGAGGCTGTGTTTCATGTTGGCTTCTCTGTTGTT 120
|
|
|
Db 130 ACTGTCACCTCCCATTCGATATGCAAGAGGCTGTGTTTCATGTTGGCTTCTCTGTTGTT 189
|
|
|
QY 121 GCCGCAATGAAATAGCTTATTTGGCTCACCGTAATCAATCAATCGATGCTTGAACAA 180
|
|
|
Db 190 GCCGCAATGAAATAGCTTATTTGGCTCACCGTAATCAATCAATCGATGCTTGAACAA 249
|
|
|
QY 181 GAATTAGTCGATTGCTTCCCAACACGGTTGCTATGCTGATACCATTCACCGTGGTATT 240
|
|
|
Db 250 GAATTAGTCGATTGCTTCCCAACACGGTTGCTATGCTGATACCATTCACCGTGGTATT 309
|
|
|
QY 241 GAATACATCCCAATATGTTGCTCCCAAGAAAGCTACTATCGATACGTTGCACGAGAA 300
|
|
|
Db 310 GAATACATCCCAATATGTTGCTCCCAAGAAAGCTACTATCGATACGTTGCACGAGAA 369
|
|
|
QY 301 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
|
|
|
Db 370 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 429
|
|
|
QY 361 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
|
|
|
Db 430 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 489
|
|
|
QY 421 ATTATGTCATCAAGATTTAGACGCTTCCGTCATTAATGATGCGCGCAACATCAATCAA 480
|
|
|
Db 490 ATTATGTCATCAAGATTTAGACGCTTCCGTCATTAATGATGCGCGCAACATCAATCAA 549
|
|
|
QY 481 CGCGATAATGTTTACCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTAACGCA 540
|
|
|
Db 550 CGCGATAATGTTTACCAACCAACTATCAGCTGTCAACATTTGTTGTTACAGTAACGCA 609
|
|
|
QY 541 CAAGGTGTCGATTATGATGCTACCAACAGTTGGGATACCAATTTGGGTTGATAATGTT 600
|
|
|
Db 610 CAAGGTGTCGATTATGATGCTACCAACAGTTGGGATACCAATTTGGGTTGATAATGTT 669
|
|
|
QY 601 TACGGTTATTTGCTGCCAACAATTCGATGATGATGATGATGATGATGATGATGATGATG 660
|
|
|
Db 670 TACGGTTATTTGCTGCCAACAATTCGATGATGATGATGATGATGATGATGATGATGATG 729
|
|
|
QY 661 ATTCTC 666
|
|
|
Db 730 ATTCTC 735

RESULT 5
AR080065
LOCUS AR080065 834 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968526.
ACCESSION AR080065
VERSION AR080065.1 GI:10006800
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 834)
AUTHORS Garman,R.D., Greenstein,J.L., Kuo,M.-C., Rogers,B.L., Franzen,H.M.,
Chen,X., Evans,S. and Shaked,Z.
TITLE T cell epitopes of the major allergens from Dermatophagoides (house
dust mite)
JOURNAL Patent: US 5968526-A 1 19-OCT-1999;
FEATURES Location/Qualifiers
source 1. 834
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

```

Query Match 99.8%; Score 664.4; DB 6; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 3.9e-181;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 60  
 DB 70 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 129

QY 61 ACTGTCACTCCCAATCGTATGCAAGGAGGCTGTGTTCAATGTTGGGCTTTCTCTGTT 120  
 DB 130 ACTGTCACTCCCAATCGTATGCAAGGAGGCTGTGTTCAATGTTGGGCTTTCTCTGTT 189

QY 121 GCCGCAACTCAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGCTGAACA 180  
 DB 190 GCCGCAACTCAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGCTGAACA 249

QY 181 GAATTAGTCGATTGTGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 240  
 DB 250 GAATTAGTCGATTGTGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 309

QY 241 GAATACATCCCAATGATGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 300  
 DB 310 GAATACATCCCAATGATGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 369

QY 301 CAATCATGCCGACGACCAATGCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 360  
 DB 370 CAATCATGCCGACGACCAATGCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 429

QY 361 CCACCAATGCAACAAATGCTGAAGCTTTGGCTCAACCCACACGCGTATGCGGTC 420  
 DB 430 CCACCAATGCAACAAATGCTGAAGCTTTGGCTCAACCCACACGCGTATGCGGTC 489

QY 421 ATTATGGCATCAAGATTTAGACGATTCGGTCATTATGATGGCCGCAACATCAATCAA 480  
 DB 490 ATTATGGCATCAAGATTTAGACGATTCGGTCATTATGATGGCCGCAACATCAATCAA 549

QY 481 CGGATTAATGGTTACCAACAACTATCAGCTGTCAACATTTGGTTACAGTACGCA 540  
 DB 550 CGGATTAATGGTTACCAACAACTATCAGCTGTCAACATTTGGTTACAGTACGCA 609

QY 541 CAAGGTGTCGATTATGATGCTGCAAGAACAGTTGGATACCAATGGGCTGATATGTT 600  
 DB 610 CAAGGTGTCGATTATGATGCTGCAAGAACAGTTGGATACCAATGGGCTGATATGTT 669

QY 601 TACGGTTATTTGCTGCAACATCGATTGATGATGATGATGATGATGATGATGATGATG 660  
 DB 670 TACGGTTATTTGCTGCAACATCGATTGATGATGATGATGATGATGATGATGATGATG 729

QY 661 ATTCTC 666  
 DB 730 ATTCTC 735

RESULT 6  
 I13095  
 LOCUS I13095 834 bp DNA linear PAT 26-JUL-1995  
 DEFINITION Sequence 1 from patent US 5433948.  
 ACCESSION I13095  
 VERSION I13095.1 GI:910444  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 834)  
 AUTHORS Thomas, W.R. and Chua, K.-Y.  
 TITLE Cloning and sequencing of allergens of dermatophagoides (house dust mite)  
 JOURNAL Patent: US 5433948-A 1 18-JUL-1995;  
 FEATURES Location/Qualifiers  
 source 1..834  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

Query Match 99.8%; Score 664.4; DB 6; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 3.9e-181;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 60  
 DB 70 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 129

QY 61 ACTGTCACTCCCAATCGTATGCAAGGAGGCTGTGTTCAATGTTGGGCTTTCTCTGTT 120  
 DB 130 ACTGTCACTCCCAATCGTATGCAAGGAGGCTGTGTTCAATGTTGGGCTTTCTCTGTT 189

QY 121 GCCGCAACTCAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGCTGAACA 180  
 DB 190 GCCGCAACTCAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGCTGAACA 249

QY 181 GAATTAGTCGATTGTGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 240  
 DB 250 GAATTAGTCGATTGTGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 309

QY 241 GAATACATCCCAATGATGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 300  
 DB 310 GAATACATCCCAATGATGTTCCCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 369

QY 301 CAATCATGCCGACGACCAATGCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 360  
 DB 370 CAATCATGCCGACGACCAATGCAACACGGTGTGTCATGTTGATACCAATCCACGTGTA 429

QY 361 CCACCAATGCAACAAATGCTGAAGCTTTGGCTCAACCCACACGCGTATGCGGTC 420  
 DB 430 CCACCAATGCAACAAATGCTGAAGCTTTGGCTCAACCCACACGCGTATGCGGTC 489

QY 421 ATTATGGCATCAAGATTTAGACGATTCGGTCATTATGATGGCCGCAACATCAATCAA 480  
 DB 490 ATTATGGCATCAAGATTTAGACGATTCGGTCATTATGATGGCCGCAACATCAATCAA 549

QY 481 CGGATTAATGGTTACCAACAACTATCAGCTGTCAACATTTGGTTACAGTACGCA 540  
 DB 550 CGGATTAATGGTTACCAACAACTATCAGCTGTCAACATTTGGTTACAGTACGCA 609

QY 541 CAAGGTGTCGATTATGATGCTGCAAGAACAGTTGGATACCAATGGGCTGATATGTT 600  
 DB 610 CAAGGTGTCGATTATGATGCTGCAAGAACAGTTGGATACCAATGGGCTGATATGTT 669

QY 601 TACGGTTATTTGCTGCAACATCGATTGATGATGATGATGATGATGATGATGATGATG 660  
 DB 670 TACGGTTATTTGCTGCAACATCGATTGATGATGATGATGATGATGATGATGATGATG 729

QY 661 ATTCTC 666  
 DB 730 ATTCTC 735

RESULT 7  
 I25468  
 LOCUS I25468 834 bp DNA linear PAT 07-OCT-1996  
 DEFINITION Sequence 1 from patent US 5552142.  
 ACCESSION I25468  
 VERSION I25468.1 GI:1605338  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 834)  
 AUTHORS Thomas, W.R. and Chua, K.-Y.  
 TITLE Cloning and sequencing of allergens of dermatophagoides (house dust mite)  
 JOURNAL Patent: US 5552142-A 1 03-SEP-1996;  
 FEATURES Location/Qualifiers  
 source 1..834  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 99.8%; Score 664.4; DB 6; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 3.9e-181;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAAGCGCTGAGTATCAATGGAAATGCTCAGCTGAAATCGATTTGGCGACAAATGCGA 60  
 DB 70 ACTAAGCGCTGAGTATCAATGGAAATGCTCAGCTGAAATCGATTTGGCGACAAATGCGA 129  
 QY 61 ACTGTCACCTCCCATTCGATGCAAGGAGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 120  
 DB 130 ACTGTCACCTCCCATTCGATGCAAGGAGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 189  
 QY 121 GCGGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATCGATTTGGCGACAA 180  
 DB 190 GCGGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATCGATTTGGCGACAA 249  
 QY 181 GAATTAGTCGATTTGCTTCCCAACACGGTTGCTGATGATACCAATTCCTGCTGATTTAC 240  
 DB 250 GAATTAGTCGATTTGCTTCCCAACACGGTTGCTGATGATACCAATTCCTGCTGATTTAC 309  
 QY 241 GAATACATCAACATCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 300  
 DB 310 GAATACATCAACATCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 369  
 QY 301 CAATCATGCGACCAACCAATGCAACAGTTTGGCTCAAAACCCACAGCGCTATTGGCGTTC 360  
 DB 370 CAATCATGCGACCAACCAATGCAACAGTTTGGCTCAAAACCCACAGCGCTATTGGCGTTC 429  
 QY 361 CCACCAATGCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 420  
 DB 430 CCACCAATGCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 489  
 QY 421 ATTATGCGATCAAGATTTAGACGCTTCCGATGATGATGATGATGATGATGATGATGAT 480  
 DB 490 ATTATGCGATCAAGATTTAGACGCTTCCGATGATGATGATGATGATGATGATGATGAT 549  
 QY 481 CCGGATATGCTTACCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAG 540  
 DB 550 CCGGATATGCTTACCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAG 609  
 QY 541 CAAGGTGCTGATTTGGATGCTGCAAGAGCTTGGGATACCAATTCCTGCTGATGATGATGAT 600  
 DB 610 CAAGGTGCTGATTTGGATGCTGCAAGAGCTTGGGATACCAATTCCTGCTGATGATGATGAT 669  
 QY 601 TAGGTTATTTGCTGCCAATCATCGATTTGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB 670 TAGGTTATTTGCTGCCAATCATCGATTTGATGATGATGATGATGATGATGATGATGATGAT 729

661 ATTCTC 666  
 730 ATTCTC 735

RESULT 8  
 BD001750  
 LOCUS Tick allergen.  
 DEFINITION BD001750  
 ACCESSION BD001750  
 VERSION BD001750.1 GI:18626309  
 KEYWORDS JP 2000078993-A/1.  
 SOURCE unclassified  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 857)  
 AUTHORS Thomas, W.R., Stewart, J.A., Turner, K.J., and Simpson, R.J.  
 TITLE Tick allergen  
 JOURNAL Patent: JP 2000078993-A 1 21-MAR-2000;  
 INST FOR CHILD HEALTH RESEARCH  
 OS Dermatophagoides sp.  
 PN JP 2000078993-A/1  
 PD 21-MAR-2000

PF 25-AUG-1999 JP 1999238264  
 PR 17-JUN-1987 AU PI2523  
 PI WAIN ROBERT THOMAS, JEFFREY ALEXANDER STEWART, KEVUN JAMES  
 TURNER,  
 PL RICHARD JOHN SIMPSON  
 PC C12N15/09 A61K39/00 A61K39/35 A61P37/08 C07K14/435//C12P21/02,  
 PC (C12P21/02, C12R1:19), C12N15/00  
 CC  
 FH Key Location/Qualifiers  
 FT CDS (1)..(735)  
 FT mat peptide (70).  
 FT Location/Qualifiers  
 FEATURES  
 source 1..857  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

ORIGIN

Query Match 99.8%; Score 664.4; DB 6; Length 857;  
 Best Local Similarity 99.8%; Pred. No. 3.9e-181;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAAGCGCTGAGTATCAATGGAAATGCTCAGCTGAAATCGATTTGGCGACAAATGCGA 60  
 DB 70 ACTAAGCGCTGAGTATCAATGGAAATGCTCAGCTGAAATCGATTTGGCGACAAATGCGA 129  
 QY 61 ACTGTCACCTCCCATTCGATGCAAGGAGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 120  
 DB 130 ACTGTCACCTCCCATTCGATGCAAGGAGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 189  
 QY 121 GCGGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATCGATTTGGCGACAA 180  
 DB 190 GCGGCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATCGATTTGGCGACAA 249  
 QY 181 GAATTAGTCGATTTGCTTCCCAACACGGTTGCTGATGATACCAATTCCTGCTGATTTAC 240  
 DB 250 GAATTAGTCGATTTGCTTCCCAACACGGTTGCTGATGATACCAATTCCTGCTGATTTAC 309  
 QY 241 GAATACATCAACATCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 300  
 DB 310 GAATACATCAACATCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 369  
 QY 301 CAATCATGCGACCAACCAATGCAACAGTTTGGCTCAAAACCCACAGCGCTATTGGCGTTC 360  
 DB 370 CAATCATGCGACCAACCAATGCAACAGTTTGGCTCAAAACCCACAGCGCTATTGGCGTTC 429  
 QY 361 CCACCAATGCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 420  
 DB 430 CCACCAATGCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAGAA 489  
 QY 421 ATTATGCGATCAAGATTTAGACGCTTCCGATGATGATGATGATGATGATGATGATGATGAT 480  
 DB 490 ATTATGCGATCAAGATTTAGACGCTTCCGATGATGATGATGATGATGATGATGATGATGAT 549  
 QY 481 CCGGATATGCTTACCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAG 540  
 DB 550 CCGGATATGCTTACCAACCAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCGACGAG 609  
 QY 541 CAAGGTGCTGATTTGGATGCTGCAAGAGCTTGGGATACCAATTCCTGCTGATGATGATGATGAT 600  
 DB 610 CAAGGTGCTGATTTGGATGCTGCAAGAGCTTGGGATACCAATTCCTGCTGATGATGATGATGAT 669  
 QY 601 TAGGTTATTTGCTGCCAATCATCGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB 670 TAGGTTATTTGCTGCCAATCATCGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 729

661 ATTCTC 666  
 730 ATTCTC 735

RESULT 9  
 ARO13797

LOCUS AR013797 1172 bp DNA linear PAT 05-DEC-1998  
DEFINITION Sequence 9 from patent US 5773002.  
ACCESSION AR013797  
VERSION AR013797.1 GI:3971251  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1172)  
AUTHORS Thomas, W.R. and Chua, K.-Y.  
TITLE Cloning and sequencing of allergens of dermatophagoides (house dust mite)  
JOURNAL Patent: US 5773002-A 9 30 JUN-1998;  
FEATURES Location/Qualifiers  
source 1..1172  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 99.8%; Score 664.4; DB 6; Length 1172;  
Best Local Similarity 99.8%; Pred. No. 4e-181;  
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60  
Db 344 ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 403  
QY 61 ACTGTCACTCCATTCGATGCAAGGAGGCTGCTTCATGTTGGGCTTCTCTGCTGTT 120  
Db 404 ACTGTCACTCCATTCGATGCAAGGAGGCTGCTTCATGTTGGGCTTCTCTGCTGTT 463  
QY 121 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTGGATCTTGCTGAACAA 180  
Db 464 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTGGATCTTGCTGAACAA 523  
QY 181 GAATTAGTCGATGCTGCTCCACACGCTGCTGATGATGATGATGATGATGATGATGATGAT 240  
Db 524 GAATTAGTCGATGCTGCTCCACACGCTGCTGATGATGATGATGATGATGATGATGATGAT 583  
QY 241 GAATACATCAACATATGCTGCTCCAGAAAGCTACTATCGATACGCTTGACAGAGAA 300  
Db 584 GAATACATCAACATATGCTGCTCCAGAAAGCTACTATCGATACGCTTGACAGAGAA 643  
QY 301 CAATCATGCGGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 644 CAATCATGCGGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 703  
QY 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTCGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
Db 704 CCACCAATGCAACCAAAATTCGTGAAGCTTTCGGCTCAAAACCCACAGCGCTATTGCCGTC 763  
QY 421 ATTATTGGCATCAAAAGATTAGACGATTCGGTATTCGATGATGATGATGATGATGATGATGAT 480  
Db 764 ATTATTGGCATCAAAAGATTAGACGATTCGGTATTCGATGATGATGATGATGATGATGATGAT 823  
QY 481 CGCGATATGTTTACCAACCAAACTATCAGCTGTCACCAATGTTGGTTACAGTACGCA 540  
Db 824 CGCGATATGTTTACCAACCAAACTATCAGCTGTCACCAATGTTGGTTACAGTACGCA 883  
QY 541 CAAGGTGTCGATTTGGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGAT 600  
Db 884 CAAGGTGTCGATTTGGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGAT 943  
QY 601 TACGGTTATTTGCTGCAACATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db 944 TACGGTTATTTGCTGCAACATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1003  
QY 661 ATTCTC 666  
Db 1004 ATTCTC 1009

LOCUS I13099 1172 bp DNA linear PAT 26-JUL-1995  
DEFINITION Sequence 9 from patent US 5433948.  
ACCESSION I13099  
VERSION I13099.1 GI:910448  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1172)  
AUTHORS Thomas, W.R. and Chua, K.-Y.  
TITLE Cloning and sequencing of allergens of dermatophagoides (house dust mite)  
JOURNAL Patent: US 5433948-A 9 18 JUL-1995;  
FEATURES Location/Qualifiers  
source 1..1172  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 99.8%; Score 664.4; DB 6; Length 1172;  
Best Local Similarity 99.8%; Pred. No. 4e-181;  
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60  
Db 344 ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 403  
QY 61 ACTGTCACTCCATTCGATGCAAGGAGGCTGCTTCATGTTGGGCTTCTCTGCTGTT 120  
Db 404 ACTGTCACTCCATTCGATGCAAGGAGGCTGCTTCATGTTGGGCTTCTCTGCTGTT 463  
QY 121 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTGGATCTTGCTGAACAA 180  
Db 464 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTGGATCTTGCTGAACAA 523  
QY 181 GAATTAGTCGATGCTGCTCCACACGCTGCTGATGATGATGATGATGATGATGATGATGAT 240  
Db 524 GAATTAGTCGATGCTGCTCCACACGCTGCTGATGATGATGATGATGATGATGATGATGAT 583  
QY 241 GAATACATCAACATATGCTGCTCCAGAAAGCTACTATCGATACGCTTGACAGAGAA 300  
Db 584 GAATACATCAACATATGCTGCTCCAGAAAGCTACTATCGATACGCTTGACAGAGAA 643  
QY 301 CAATCATGCGGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 644 CAATCATGCGGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 703  
QY 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTCGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
Db 704 CCACCAATGCAACCAAAATTCGTGAAGCTTTCGGCTCAAAACCCACAGCGCTATTGCCGTC 763  
QY 421 ATTATTGGCATCAAAAGATTAGACGATTCGGTATTCGATGATGATGATGATGATGATGATGAT 480  
Db 764 ATTATTGGCATCAAAAGATTAGACGATTCGGTATTCGATGATGATGATGATGATGATGATGAT 823  
QY 481 CGCGATATGTTTACCAACCAAACTATCAGCTGTCACCAATGTTGGTTACAGTACGCA 540  
Db 824 CGCGATATGTTTACCAACCAAACTATCAGCTGTCACCAATGTTGGTTACAGTACGCA 883  
QY 541 CAAGGTGTCGATTTGGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGAT 600  
Db 884 CAAGGTGTCGATTTGGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATATGAT 943  
QY 601 TACGGTTATTTGCTGCAACATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Db 944 TACGGTTATTTGCTGCAACATGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1003  
QY 661 ATTCTC 666  
Db 1004 ATTCTC 1009





Db	481	CCGGTAATAGTTACCAACCAAACTATCACGCTGTGCAACATTGTTGGTTACAGTAACGCA	540
Qy	541	CAAGGTGCGATTATTGGATCGTACGAAACAGTTGGGATACCAATTTGGGGTGATTAATGGT	600
Db	541	CAAGGTGCGATTATTGGATCGTACGAAACAGTTGGGATACCAATTTGGGGTGATTAATGGT	600
Qy	601	TACGTTATTTTTGTCGCCAACATCGATTGTGATGATTGAAGAATATCCATATGTTGTC	660
Db	601	TACGTTATTTTTGTCGCCAACATCGATTGTGATGATTGAAGAATATCCATATGTTGTC	660
Qy	661	ATTCTC	666
Db	661	ATTCTC	666

RESULT 13	AX1114208	LOCUS	AX1114208	669 bp	DNA	linear	PAT 11-MAY-2001
DEFINITION	Sequence 16 from Patent WO0129078.						
ACCESSION	AX1114208						
VERSION	AX1114208.1						
KEYWORDS	GI:14031184						
SOURCE	Dermatophagoides pteronyssinus (European house dust mite)						
ORGANISM	Dermatophagoides pteronyssinus						
	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;						
	Acariiformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;						
	Pyroglyphidae; Dermatophagoides.						
REFERENCE	1						

ORIGIN	YWIL"	Query Match	99.3%;	Score 661.2;	DB 6;	Length 669;
		Best Local Similarity	99.5%;	Pred. No. 3.2e-180;		
		Matches 663;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	ACTAACGCGCTGCAGTATCAATGAAATGCTCCAGCTGAAATCGAATTCGCAAAATCGGA	60			
Db	4	ACTAACGCGCTGCAGTATCAATGAAATGCTCCAGCTGAAATCGAATTCGCAAAATCGGA	63			
QY	61	ACTGTCACTCCCATTCGTATGCAAGAGAGCGTGTGGTTCATGTGGGCTTTCTCTGTGTT	120			
Db	64	ACTGTCACTCCCATTCGTATGCAAGAGAGCGTGTGGTTCATGTGGGCTTTCTCTGTGTT	123			
QY	121	GCCGCAACTGAATCAGCTTAATTTGGCTCACCGTCTACCGTATCAATCATTTGGATCTTGCTGAACAA	180			
Db	124	GCCGCAACTGAATCAGCTTAATTTGGCTCACCGTATCAATCATTTGGATCTTGCTGAACAA	183			
QY	181	GAATTAGTCGATTGTGCTTCCCAACACGGTTGTGATGGTGATACCATTCACAGCTGGTATT	240			
Db	184	GAATTAGTCGATTGTGCTTCCCAACACGGTTGTGATGGTGATACCATTCACAGCTGGTATT	243			
QY	241	GAATACATCCAACTAATGTGTGCTTCAAGAAAGCTACTATCGATACGTTGCACGAA	300			
Db	244	GAATACATCCAACTAATGTGTGCTTCAAGAAAGCTACTATCGATACGTTGCACGAA	303			

Qy	301	CAATCATGCCGACACACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCGAAATTTAC	360
Db	304	CAATCATGCCGACGACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCGAAATTTAC	363
Qy	361	CCACCAAAATGCAAAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACACGGCTATTGCCGTC	420
Db	364	CCACCAAAATGCAAAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACACGGCTATTGCCGTC	423
Qy	421	ATTATTGGCATCAAAAGATTTAGACGGCATTCGGTCAATTATGATGGCCGAAACAATTCATCAA	480
Db	424	ATTATTGGCATCAAAAGATTTAGACGGCATTCGGTCAATTATGATGGCCGAAACAATTCATCAA	483
Qy	481	CGCGATTAATGGTTACCAACCAAAACTATCACGGCTGTCAACATTTGTTGTTACAGTAACGCA	540
Db	484	CGCGATTAATGGTTACCAACCAAAACTATCACGGCTGTCAACATTTGTTGTTACAGTAACGCA	543
Qy	541	CAGAGTGTCGATTAATGGATTCGTACGAAACAGTTGGGATACCAATTCGGGTGATAATGGT	600
Db	544	CAGAGTGTCGATTAATGGATTCGTACGAAACAGTTGGGATACCAATTCGGGTGATAATGGT	603
Qy	601	TACGGTTATTTTCTGCCAAACATCGATTTGATGATGATTGAAGAATATCCATATGTTGTC	660
Db	604	TACGGTTATTTTCTGCCAAACATCGATTTGATGATGATTGAAGAATATCCATATGTTGTC	663
Qy	661	ATTCTC 666	
Db	664	ATTCTC 669	
RESULT 14			
AX114210/c			
LOCUS	AX114210	669 bp	DNA linear PAT 11-MAY-2001
DEFINITION	Sequence 18 from Patent WO0129078.		
ACCESSION	AX114210		
VERSION	AX114210.1	GI:14031186	
KEYWORDS	Dermatophagoides pteronyssinus (European house dust mite)		
SOURCE	Dermatophagoides pteronyssinus		
ORGANISM	Dermatophagoides pteronyssinus		
	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;		
	Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;		
	Pyroglyphidae; Dermatophagoides.		
REFERENCE	1		
AUTHORS	Best, E.A. and McDermott, M.J.		
TITLE	Method for production and use of mite group 1 proteins		
JOURNAL	Patent: WO 0129078-A 18 26-APR-2001;		

FEATURES	source	Location/Qualifiers
	1..669	
		/organism="Dermatophagoides pteronyssinus"
		/mol_type="unassigned DNA"
		/db_xref="taxon:6956"
ORIGIN		
	Query Match	99.3%; Score 661.2; DB 6; Length 669;
	Best Local Similarity	99.5%; Pred. No. 3.2e-180;
	Matches 663; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	1	ACTAACGCCTGCAGTATCAATGAAATGCTCCAGCTGAAATCGATTGGCACAATTCGGA 60
Db	666	ACTAACGCCTGCAGTATCAATGAAATGCTCCAGCTGAAATCGATTGGCACAATTCGGA 607
QY	61	ACTGTCACTCCCATTTTCGTATGCAAGAGAGCTGGTTCATGTTGGCTTTCCTCTGCTGTT 120
Db	606	ACTGTCACTCCCATTTTCGTATGCAAGAGAGCTGGTTCATGTTGGCTTTCCTCTGCTGTT 547
QY	121	GCGCACAACCTGAATCAGCTTATTTTGGCTCACCGTAAATCAATTCATTGGATCTTGCTGAACAA 180
Db	546	GCGCACAACCTGAATCAGCTTATTTTGGCTTACCGTAAATCAATTCATTGGATCTTGCTGAACAA 487
QY	181	GAATTAGTCGATTGTGCTTCCCAACACGGTTGTCTAGTGAATACCAATTCACACGTGTGATT 240
Db	486	GAATTAGTCGATTGTGCTTCCCAACACGGTTGTCTAGTGAATACCAATTCACACGTGTGATT 427

QY	241	GAATACATCCAAATAATGGTGTCTCCAAAGAAAGCTACTATCGATAGCTTGCACGAGAA	300
Db	426	GAATACATCCAAATAATGGTGTCTCCAAAGAAAGCTACTATCGATAGCTTGCACGAGAA	367
QY	301	CAATCATCCGACCCAAATGCAACAAGTTTCGGTATCTCAAACTATTGCCAAATTTAC	360
Db	366	CAATCATCCGACCCAAATGCAACAAGTTTCGGTATCTCAAACTATTGCCAAATTTAC	307
QY	361	CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC	420
Db	306	CCACCAATGTAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC	247
QY	421	ATTATGCGCATCAAGATTTAGACGCAATTCGGTCATTTATGATGCGCGAAACAATCAATCAA	480
Db	246	ATTATGCGCATCAAGATTTAGACGCAATTCGGTCATTTATGATGCGCGAAACAATCAATCAA	187
QY	481	CGCGATAATGGTTACCAACCAAACTATCACGCTGTCAACATTTGTTGGTTACAGTAAGCA	540
Db	186	CGCGATAATGGTTACCAACCAAACTATCACGCTGTCAACATTTGTTGGTTACAGTAAGCA	127
QY	541	CAAGGTGTCGATTTATGATCGTACGAAACAGTTGGGATACCAATTTGGGGTGATAATGGT	600
Db	126	CAAGGTGTCGATTTATGATCGTACGAAACAGTTGGGATACCAATTTGGGGTGATAATGGT	67
QY	601	TACGGTTATTTTGGTGCACCAATCGATTTGATGATGATTGAAGAAATATCCATATGTTGTC	660
Db	66	TACGGTTATTTTGGTGCACCAATCGATTTGATGATGATTGAAGAAATATCCATATGTTGTC	7
QY	661	ATTCTC 666	
Db	6	ATTCTC 1	
RESULT 15			
LOCUS	CO786568	849 bp	DNA linear PAT 24-MAR-2004
DEFINITION	Sequence 5 from Patent WO2004019978.		
ACCESSION	CO786568		
VERSION	CO786568.1 GI:45721609		
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Chua,K.Y. and Liew,L.N.		
TITLE	Recombinant nucleic acid useful for inducing protective immune response against allergens		
JOURNAL	Patent: WO 2004019978-A 5 11-MAR-2004; NATIONAL UNIVERSITY OF SINGAPORE (SG)		
FEATURES	Location/Qualifiers		
source	1..849		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="chimeric gene that encodes the Mus musculus LAMP-1 leader sequence, the e n tire Der p 1 gene product and the Mus musculus LAMP-1 transmembrane and cytoplasmic domain"		
ORIGIN			
Query Match	99.3%;	Score 661.2;	DB 6; Length 849;
Best Local Similarity	99.5%;	Pred.No. 3.3e-180;	
Matches 663;	Conservative 0;	Mismatches 3;	Indels 0; Gaps 0;
QY	1	ACTAAGCGCTCGAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGCGACAAATGCGA	60
Db	70	ACTAAGCGCTCGAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGCGACAAATGCGA	129
QY	61	ACTGTCACTCCCATTCGATGCAAGAGCGTGTGGTTTCATGTTGGGCTTCTCTGGTGT	120
Db	130	ACTGTCACTCCCATTCGATGCAAGAGCGTGTGGTTTCATGTTGGGCTTCTCTGGTGT	189
QY	121	GCCGCAACTGATCAGCTTATTGGCTCACCGTAAATCAATGATCTTGGTGAACAA	180

Db	190	GCGCAACTGAATCAGCTTATTGGCTTACCGTAAATCAATCATTTGGATCTTGCTGAACAA	249
QY	181	GAATTAAGTGTGATGCTTCCCAACACGCTTGTTCATGGTGATACCAATTCACGTGGTATT	240
Db	250	GAATTAAGTGTGATGCTTCCCAACACGCTTGTTCATGGTGATACCAATTCACGTGGTATT	309
QY	241	GAATATACATCAACATTAATGGTGTCTGTCGAAGAAAGCTACTATCGATACGTTGCACGAGAA	300
Db	310	GAATATACATCAACATTAATGGTGTCTGTCGAAGAAAGCTACTATCGATACGTTGCACGAGAA	369
QY	301	CAATCATGCGACCCCAAAATGTCRCAAGTTCGCTATCTCAAACTATTGCCAATTTAC	360
Db	370	CAATCATGCGACCCCAAAATGTCRCAAGTTCGCTATCTCAAACTATTGCCAATTTAC	429
QY	361	CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC	420
Db	430	CCACCAATGTAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC	489
QY	421	ATTATGCGCATCAAGATTTAGACCCATTCGCTCATATTATGATGGCCGAAACAATCAATCAA	480
Db	490	ATTATGCGCATCAAGATTTAGACCCATTCGCTCATATTATGATGGCCGAAACAATCAATCAA	549
QY	481	CGCGATAATGGTTACCAACCAAACTATCACGCTGTCAACATTTGTTGGTTACAGTAAGCA	540
Db	550	CGCGATAATGGTTACCAACCAAACTATCACGCTGTCAACATTTGTTGGTTACAGTAAGCA	609
QY	541	CAAGGTGTCGATTTATGATCGTACGAAACAGTGGGATACCAATTTGGGGTGATAATGGT	600
Db	610	CAAGGTGTCGATTTATGATCGTACGAAACAGTGGGATACCAATTTGGGGTGATAATGGT	669
QY	601	TACGGTTATTTTGGCGCAACATCGATTTGATGATGATTGAAGATATCCATATGTTGTC	660
Db	670	TACGGTTATTTTGGCGCAACATCGATTTGATGATGATTGAAGATATCCATATGTTGTC	729
QY	661	ATTCTC 666	
Db	730	ATTCTC 735	

Search completed: October 28, 2004, 23:43:56  
Job time : 3508.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 20:36:09 ; Search time 6942.1 seconds  
(without alignments)  
3495.894 Million cell updates/sec

Title: US-09-867-159A-1  
Perfect score: 666  
Sequence: 1 actaacgctgcagctatcaa.....atccatattgttcattctc 666

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsel:\*  
9: gb\_gsel2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	139	20.9	679	5	BQ834765	BQ834765 Po.ad_03B
2	138.6	20.8	708	5	BQ834626	BQ834626 Po.ad_01D
3	137.2	20.6	686	5	BQ834896	BQ834896 Po.ad_02D
4	137	20.6	652	5	BQ834843	BQ834843 Po.ad_04B
5	136.8	20.5	686	5	BQ834615	BQ834615 Po.ad_01B
6	136.8	20.5	704	5	BQ834782	BQ834782 Po.ad_03D
7	112	16.8	638	5	BQ834852	BQ834852 Po.ad_04C
8	111.6	16.8	612	5	BQ835052	BQ835052 Po.ad_06F
9	102.4	15.4	509	6	CA305277	CA305277 ESSU900 S
10	87.4	13.1	557	5	BQ834840	BQ834840 Po.ad_04A
11	71.4	10.7	953	5	BQ792027	BQ792027 SJF2AGF02
12	71.2	10.7	740	4	BG583083	BG583083 EST484833
13	70.4	10.6	661	6	CA396057	CA396057 rg0906.Y
14	70	10.5	420	5	BQ090548	BQ090548 rc66h04.Y
15	69.2	10.4	734	5	BQ716792	BQ716792 SJM2AXB11
C 16	69	10.4	721	4	BQ401582	BQ401582 BJ401582
C 17	68	10.2	659	4	BQ339781	BQ339781 BJ339781
C 18	67.8	10.2	508	6	CA524770	CA524770 KS12043A0
C 19	67.8	10.2	734	4	BQ432452	BQ432452 BJ432452
C 20	67.8	10.2	735	4	BQ436222	BQ436222 BJ436222
C 21	67.6	10.2	688	7	CF673480	CF673480 tac98f06
C 22	66.6	10.0	622	6	CF181972	CF181972 ISO4D8F.I
C 23	66.4	10.0	721	4	BQ402552	BQ402552 BJ402552
C 24	66.4	10.0	748	5	BQ474518	BQ474518 biphyllus

25	66.4	10.0	860	5	BU719074	BU719074 SJM2QD003
C 26	66.2	9.9	734	1	AU214559	AU214559 AU214559
C 27	66	9.9	822	5	BU795793	BU795793 SJF2DYCO7
28	65.8	9.9	711	5	BU798159	BU798159 SJF2AAH05
C 29	65.8	9.9	719	4	BJ402553	BJ402553 BJ402553
30	65.8	9.9	738	5	BU716851	BU716851 SJM2AXG10
31	65.8	9.9	755	5	BU714644	BU714644 SJMBOF01
32	65.6	9.8	633	7	CK426618	CK426618 rx18a04.Y
33	65.4	9.8	654	6	CA753822	CA753822 BR0400030
C 34	65.4	9.8	707	5	BO474510	BO474510 biphyllus
C 35	65.4	9.8	736	4	BJ344513	BJ344513 BJ344513
C 36	65.2	9.8	719	4	BU372531	BU372531 BU372531
37	65	9.8	807	7	CO162301	CO162301 FLD1_34_A
38	64.8	9.7	729	5	BU715806	BU715806 SJM2ABG06
39	64.6	9.7	812	5	BU715970	BU715970 SJM2AJE04
40	64.6	9.7	869	5	BU793212	BU793212 SJF2BEC05
C 41	64.2	9.6	607	4	BJ345832	BJ345832 BJ345832
42	64.2	9.6	684	5	BU715911	BU715911 SJM2ADH04
43	64	9.6	1044	3	CR688981	CR688981 tetradon
C 44	63.8	9.6	681	7	CF779911	CF779911 tad05b08
C 45	63.6	9.5	676	4	BJ436012	BJ436012 BJ436012

ALIGNMENTS

BQ834765 679 bp mRNA linear EST 08-AUG-2002  
Po.ad\_03B08 TEXT1 Psoroptes ovis mixed Psoroptes ovis cDNA clone  
Po.ad\_03B08'5, similar to CAB58493 SEQUENCE 5 FROM PATENT  
WO9424281. unidentified, mRNA sequence.

ACCESSION BQ834765  
VERSION BQ834765.1  
KEYWORDS GI:22139079  
SOURCE Psoroptes ovis (sheep scab mite)  
ORGANISM Psoroptes ovis

REFERENCE 1 (bases 1 to 679)  
AUTHORS Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and Knox,D.P.  
TITLE Expressed sequence tag survey of gene expression in the scab mite Psoroptes ovis-allergens, proteases and free-radical scavengers  
JOURNAL Parasitology 126 (Pt 5), 451-460 (2003)  
MEDLINE 22677800  
PUBMED 12793649

COMMENT Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared for Dr David Knox, Moredun Research Institute, UK Sequencing was performed by Mark Welsh, ICAPB, Edinburgh  
PCR Primers  
FORWARD: M13R  
BACKWARD: T7PL  
Plate: 03 row: B column: 08  
High quality sequence stop: 542.

FEATURES  
source /organism="Psoroptes ovis"  
mol\_type="mRNA"  
/db\_xref="taxon:83912"  
/clone="Po.ad\_03B08"  
/sex="Mixed"  
/dev\_stage="Mixed (predominantly adult)"  
/clone\_lib="Psoroptes ovis mixed"

/notes="Vector: Lambda Triplex2; Site 1: Sfi1A; Site 2: Sfi1B; Psoroptes ovis is the infective agent causing sheep scab, a serious dermatitis in sheep. The cDNA was obtained from mites harvested from the skin of previously unexposed sheep harbouring a moderate challenge infection. The cDNA was cloned using the Clontech Lambda Triplex2 Smart cDNA cloning system."

## ORIGIN

Query Match 20.9%; Score 139; DB 5; Length 679;  
 Best Local Similarity 69.1%; Pred. No. 1.2e-31;  
 Matches 190; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 22 GGAATGCTCCAGCTGAAATCGAATTTGCGACAATCGGAACCTGTCACCTCCCTTCGTATG 81  
 DB 405 GTAACATTCCTCAATGAAATGATTTACGCTGCTTTGGTTATGTAACAAAATCAAGAT 464

QY 82 CAAGAGGCTGTGGTTCATGTTGGCTTTCTCTGCTGTTGCCGCAACTGAATCAGCTTAT 141  
 DB 465 CAAGTTGCCCTGTGGTTCATGCTGGCTTTCTCTGCTGTTGCCGCAACTGAATCAGCTTAT 524

QY 142 TTGGCTCACCGTAAATCAATCATTTGGATTTGCTGCAACAGAAATTTAGTTCGATTGCTTCC 201  
 DB 525 TTATCATCATATGATATCATTTAGATCTTTCTGACACAGAAATTTGCTGACTGCTTCA 584

QY 202 CAACACGGTTGTCATGCTGATACATTCACGCTGCTGATGATATGATATGATATGATG 261  
 DB 585 CAACACGGTTGCTGCTGATACATTTCTAAACGGTTTACGTTATATTTCAAAAGATG 644

QY 262 GTGCTCAACAAAGCTACTATCATGCTGTTGACG 296  
 DB 645 GTGCTCAACAAAGCTACTATCATGCTGTTGACG 679

## RESULT 2

BQ834626 708 bp mRNA linear EST 08-AUG-2002  
 LOCUS  
 DEFINITION  
 PO ad 01D02 TEXF1 Psoroptes ovis mixed Psoroptes ovis cDNA clone  
 PO ad 01D02 5, similar to CAB58493 SEQUENCE 5 FROM PATENT  
 WO9424281. unidentified, mRNA sequence.

## ACCESSION

VERSION  
 BQ834626  
 EST.

## KEYWORDS

SOURCE

## ORGANISM

Psoroptes ovis (sheep scab mite)  
 Psoroptes ovis  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 Acariformes; Sarcopiformes; Astigmata; Psoroptidia; Sarcopitoidea;  
 Psoroptidae; Psoroptes.

## REFERENCE

AUTHORS  
 Kenyon, F., Welsh, M., Parkinson, J., Whitton, C., Blaxter, M.L. and Knox, D.P.  
 TITLE  
 Expressed sequence tag survey of gene expression in the scab mite  
 Psoroptes ovis--allergens, proteases and free-radical scavengers  
 JOURNAL  
 Parasitology 126 (Pt 5), 451-460 (2003)  
 MEDLINE  
 22677800  
 PUBMED  
 12793649

## COMMENT

Contact: Blaxter ML  
 Institute of Cell, Animal and Population Biology  
 University of Edinburgh  
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
 3JT, UK.  
 Tel: +44 131 650 6760  
 Fax: +44 131 670 5450  
 Email: mark.blaxter@ed.ac.uk

The library was prepared for Dr David Knox, Moredun Research  
 Institute, UK Sequencing was performed by Mark Welsh, ICAPB,  
 Edinburgh  
 PCR Primers  
 FORWARD: M13R  
 BACKWARD: T7PL  
 Plate: 01 row: D column: 02  
 High quality sequence stop: 540.

Location/Qualifiers

## FEATURES

## source

1. .708  
 /organism="Psoroptes ovis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:83912"  
 /clone="Po ad 01D02"  
 /sex="Mixed"  
 /dev\_stage="Mixed (predominantly adult)"  
 /clone\_lib="Psoroptes ovis mixed"  
 /note="Vector: Lambda Triplex2; Site 1: Sfi1A; Site 2: Sfi1B; psoroptes ovis is the infective agent causing sheep scab, a serious dermatitis in sheep. The cDNA was obtained from mites harvested from the skin of previously unexposed sheep harbouring a moderate challenge infection. The cDNA was cloned using the Clontech Lambda Triplex2 Smart cDNA cloning system."

## ORIGIN

Query Match 20.8%; Score 138.6; DB 5; Length 708;  
 Best Local Similarity 65.9%; Pred. No. 1.6e-31;  
 Matches 201; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 25 AATGCTCAGCTGAAATCGAATTTGCGACAATCGGAACCTGTCACCTCCCTTCGTATGCAA 84  
 DB 402 AATTTACCACACGAAATTTAGTTACGCTGCTATGGTCATGTCACAAAATCAAGATCAA 461

QY 85 GGAGGCTGTGGTTCATGTTGGCTTTCTCTGCTGTTGCCGCAACTGAATCAGCTTATTTG 144  
 DB 462 GCTCATGTTGGCTCATGCTGGCTTTCTCAGTGTGCTACAGTCGAATCAACATCTTA 521

QY 145 GCTCACCGTAAATCAATCATTTGGATTTGCTGCAACAGAAATTTAGTTCGATTGCTTCCCAA 204  
 DB 522 GCATACCGTAATGATATCATTTAGATCTTTCTGACACAGAAATTTAGTTCGATCACA 581

QY 205 CACGTTGTCATGCTGATACCATTCACGCTGCTGATGATATGATATGATATGATGCTC 264  
 DB 582 CACGTTGTCGACGCTGACACAAATGCCCGCTTGGTTATATTCAGAAAATGCTGT 641

QY 265 GTCCAAAGAACTACTATCATGCTGCTGACGAGAAACAATCATGCCGACCAACCAATGCA 324  
 DB 642 GTTCAAGAACGAATTTATCCATCCAGCTCGCGCGGAGATGTCAGAGGTTGATAGT 701

QY 325 CAACG 329  
 DB 702 CGACG 706

## RESULT 3

BQ834696 686 bp mRNA linear EST 08-AUG-2002  
 LOCUS  
 DEFINITION  
 PO ad 02D04 TEXF1 Psoroptes ovis mixed Psoroptes ovis cDNA clone  
 PO ad 02D04 5, similar to CAB58493 SEQUENCE 5 FROM PATENT  
 WO9424281. unidentified, mRNA sequence.

## ACCESSION

VERSION  
 BQ834696  
 EST.

## KEYWORDS

SOURCE

## ORGANISM

Psoroptes ovis (sheep scab mite)  
 Psoroptes ovis  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 Acariformes; Sarcopiformes; Astigmata; Psoroptidia; Sarcopitoidea;  
 Psoroptidae; Psoroptes.

## REFERENCE

AUTHORS  
 Kenyon, F., Welsh, M., Parkinson, J., Whitton, C., Blaxter, M.L. and Knox, D.P.  
 TITLE  
 Expressed sequence tag survey of gene expression in the scab mite  
 Psoroptes ovis--allergens, proteases and free-radical scavengers  
 JOURNAL  
 Parasitology 126 (Pt 5), 451-460 (2003)  
 MEDLINE  
 22677800  
 PUBMED  
 12793649

## COMMENT

Contact: Blaxter ML  
 Institute of Cell, Animal and Population Biology  
 University of Edinburgh  
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
 3JT, UK.

Tel: +44 131 650 6760  
 Fax: +44 131 670 5450  
 Email: mark.blaxter@ed.ac.uk  
 The library was prepared for Dr David Knox, Moredun Research Institute, UK Sequencing was performed by Mark Welsh, ICAPB, Edinburgh  
 PCR Primers  
 FORWARD: M13R  
 BACKWARD: T7PL  
 Plate: 02 row: D column: 04  
 High quality sequence stop: 534.  
 Location/Qualifiers  
 1. 686  
 /organism="Psoroptes ovis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:83912"  
 /clone="Po ad 02D04"  
 /sex="Mixed"  
 /dev\_stage="Mixed (predominantly adult)"  
 /clone\_lib="Psoroptes ovis mixed"  
 /note="Vector: Lambda Triplex2, Site 1: SfilA; Site 2: SfilB; Psoroptes ovis is the infective agent causing sheep scab, a serious dermatitis in sheep. The cDNA was obtained from mites harvested from the skin of previously unexposed sheep harbouring a moderate challenge infection. The cDNA was cloned using the Clontech Lambda Triplex2 Smart cDNA cloning system."  
 FEATURES  
 source  
 1. 686  
 /organism="Psoroptes ovis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:83912"  
 /clone="Po ad 02D04"  
 /sex="Mixed"  
 /dev\_stage="Mixed (predominantly adult)"  
 /clone\_lib="Psoroptes ovis mixed"  
 /note="Vector: Lambda Triplex2, Site 1: SfilA; Site 2: SfilB; Psoroptes ovis is the infective agent causing sheep scab, a serious dermatitis in sheep. The cDNA was obtained from mites harvested from the skin of previously unexposed sheep harbouring a moderate challenge infection. The cDNA was cloned using the Clontech Lambda Triplex2 Smart cDNA cloning system."  
 ORIGIN  
 Query Match 20.6%; Score 137.2; DB 5; Length 686;  
 Best Local Similarity 67.5%; Pred. No. 4.3e-31;  
 Matches 193; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
 QY 25 AATGCTCAGCTGAATCGATTTCGACAAATGCGAACTGTCACCTCCCATCGTATGCAA 84  
 Db 401 AATTTACACACGAATGATTTAGCTGTGCTGATGTCACAAATCAAGAATCAA 460  
 QY 85 GGAGGCTGTGTTCAATGTTGGGCTTTCTCTGTTGGCGCAACTGAATCAGCTTATTG 144  
 Db 461 GTGCTATGGCTCATGCTGGCTTTCTCCAGTGTGCTACAGTCGAATCAACATCTTA 520  
 QY 145 GCTCAGGTAAATCAATCATTTGATCTTCTGTAACAAAGTAATGATGCTTCCCAA 204  
 Db 521 GCATACCGTAATGATCAATTAATGATCTTTCTGAACAAAGTAATGATGCTTCCCAA 580  
 QY 205 CACGGTGTGCTATGCTATGATTCACATTCACGCTGTTGTAATGATACATCCAAATAGTGTGTC 264  
 Db 581 CACGGTGTGCTATGCTATGATTCACATTCACGCTGTTGTAATGATACATCCAAATAGTGTGTC 640  
 QY 265 GTCCAAAGAAAGCTATGCTATGATGCTTGCACGAGAACAAATCATGCC 310  
 Db 641 GTTGAAGAACGAAGTTATCCATACCGAGTCCGCGCGGAAGATGTC 686  
 RESULT 4  
 BQ834843  
 LOCUS  
 DEFINITION  
 Po ad 04B02 TEXF1 Psoroptes ovis mixed Psoroptes ovis cDNA clone  
 Po ad 04B02 5' similar to CAB58493 SEQUENCE 5 FROM PATENT  
 WO9424281. unidentified, mRNA sequence.  
 ACCESSION  
 BQ834843  
 VERSION  
 BQ834843.1 GI:22139157  
 KEYWORDS  
 EST.  
 SOURCE  
 Psoroptes ovis (sheep scab mite)  
 ORGANISM  
 Psoroptes ovis  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Sarcoptoidea;  
 Psoroptidae; Psoroptes.  
 REFERENCE  
 1 (bases 1 to 686)  
 Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and  
 Knox,D.P.  
 Expressed sequence tag survey of gene expression in the scab mite  
 Psoroptes ovis--allergens, proteases and free-radical scavengers

Parasitology 126 (Pt 5), 451-460 (2003)  
 22677800  
 MEDLINE  
 12793649  
 PUBMED  
 COMMENT  
 Contact: Blaxter ML  
 Institute of Cell, Animal and Population Biology  
 University of Edinburgh  
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK.  
 Tel: +44 131 650 6760  
 Fax: +44 131 670 5450  
 Email: mark.blaxter@ed.ac.uk  
 The library was prepared for Dr David Knox, Moredun Research Institute, UK Sequencing was performed by Mark Welsh, ICAPB, Edinburgh  
 PCR Primers  
 FORWARD: M13R  
 BACKWARD: T7PL  
 Plate: 04 row: B column: 02  
 High quality sequence stop: 490.  
 Location/Qualifiers  
 1. 652  
 /organism="Psoroptes ovis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:83912"  
 /clone="Po ad 04B02"  
 /sex="Mixed"  
 /dev\_stage="Mixed (predominantly adult)"  
 /clone\_lib="Psoroptes ovis mixed"  
 /note="Vector: Lambda Triplex2, Site 1: SfilA; Site 2: SfilB; Psoroptes ovis is the infective agent causing sheep scab, a serious dermatitis in sheep. The cDNA was obtained from mites harvested from the skin of previously unexposed sheep harbouring a moderate challenge infection. The cDNA was cloned using the Clontech Lambda Triplex2 Smart cDNA cloning system."  
 ORIGIN  
 Query Match 20.6%; Score 137; DB 5; Length 652;  
 Best Local Similarity 65.6%; Pred. No. 4.9e-31;  
 Matches 200; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
 QY 25 AATGCTCAGCTGAATCGATTTCGACAAATGCGAACTGTCACCTCCCATCGTATGCAA 84  
 Db 346 AATTTACACACGAATGATTTAGCTGTGCTGATGTCACAAATCAAGAATCAA 405  
 QY 85 GGAGGCTGTGTTCAATGTTGGGCTTTCTCTGTTGGCGCAACTGAATCAGCTTATTG 144  
 Db 406 GCTGCAATGGCTCATGCTGGGCTTTCTCCAGTGTGCTACAGTCGAATCAACATCTTA 465  
 QY 145 GCTCAGGTAAATCAATCATTTGATGCTTGCACGAGAACAAATAGTGTGCTTCCCAA 204  
 Db 466 GCATACCGTAATGATCAATTAATGATCTTTCTGAACAAAGTAATGATGCTTCCCAA 525  
 QY 205 CACGGTGTGCTATGCTATGATTCACATTCACGCTGTTGTAATGATACATCCAAATAGTGTGTC 264  
 Db 526 CACGGTGTGCTATGCTATGATTCACATTCACGCTGTTGTAATGATACATCCAAATAGTGTGTC 585  
 QY 265 GTCCAAAGAAAGCTATGCTATGATGCTTGCACGAGAACAAATAGTGTGCTTCCCAA 324  
 Db 586 GTTGAAGAACGAAGTTATCCATACCGAGTCCGCGCGGAAGATGTCAAAGAGTTGATAGT 645  
 QY 325 CAACG 329  
 Db 646 CGACG 650  
 RESULT 5  
 BQ834615  
 LOCUS  
 DEFINITION  
 Po ad 01B12 TEXF1 Psoroptes ovis mixed Psoroptes ovis cDNA clone  
 Po ad 01B12 5' similar to CAB58493 SEQUENCE 5 FROM PATENT  
 WO9424281. unidentified, mRNA sequence.  
 ACCESSION  
 BQ834615

```

VERSION B0834615.1 GI:22138929
KEYWORDS Psoroptes ovis (sheep scab mite)
SOURCE Psoroptes ovis
ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
          Acariformes; Sarcoptiformes; Astigmata; Psoroptida; Sarcoptoidea;
          Psoroptidae; Psoroptes.
REFERENCE 1 (bases 1 to 704)
AUTHORS Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and
          Knox,D.P.
TITLE Expressed sequence tag survey of gene expression in the scab mite
JOURNAL Psoroptes ovis--allergens, proteases and free-radical scavengers
MEDLINE Parasitology 126 (Pt 5), 451-460 (2003)
PUBMED 22677800
COMMENT Contact: Blaxter ML
          Institute of Cell, Animal and Population Biology
          University of Edinburgh
          Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
          3JT, UK.
          Tel: +44 131 650 6760
          Fax: +44 131 670 5450
          Email: mark.blaxter@ed.ac.uk
          The library was prepared for Dr David Knox, Moredun Research
          Institute, UK Sequencing was performed by Mark Welsh, ICAPB,
          Edinburgh
PCR Primers
FORWARD: M13R
BACKWARD: T7PL
Plate: 01 row: B column: 12
High quality sequence stop: 535.
FEATURES             source
    source
        Location/Qualifiers
            1..704
                /organism="Psoroptes ovis"
                /mol_type="mRNA"
                /db_xref="taxon:83912"
                /clone="Po ad_01B12"
                /sex="Mixed"
                /dev_stage="Mixed (predominantly adult)"
                /clone_lib="Psoroptes ovis mixed"
                /note="Vector: Lambda Triplex2; Site.1: SfilA; Site.2:
                SfilB; Psoroptes ovis is the infective agent causing sheep
                scab, a serious dermatitis in sheep. The cDNA was
                obtained from mites harvested from the skin of previously
                unexposed sheep harbouring a moderate challenge
                infection. The cDNA was cloned using the Clontech Lambda
                Triplex2 Smart cDNA cloning system."
ORIGIN
Query Match      20.6%; Score 137; DB 5; Length 704;
Best Local Similarity 65.6%; Pred. No. 5e-31;
Matches 200; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY      25  AATGCTCCACTGAATCGATTTGCGACAATCGCAACTGTCACCTCCCATTCGTATGCAA 84
DB      400  AATTACACACGAAATGATTTACGTGCTATGGGTGATGTCACAAAATCAAGATCAA 459
QY      85  GGAGGCTGTGGTTCATGTTGGGCTTTCTCTGGTGTGGCGCAACTGCAATCAGCTTATTG 144
DB      460  GCTGCATGTGGCTCATGCTGGCTTTCTCCAGTGTGCTACAGTCAATCAACATCTTA 519
QY      145  GTCACCGTAATCAATCATTTGGATCTTCTGTGAACAAGAAATAGTCATGTCGCCAA 204
DB      520  GCATATCGTAAATGATCATTTAGATCTTTCTGAACAAGAATTTAGTTGACGTGCATCACA 579
QY      205  CACGGTTGTCATGTTGATATCATTCACAGTGGTATTGAATATACATCAACAATATGTTGTC 264
DB      580  CACGGTTGCGAGGTGACAAATGCCCGGTTTGGGTTTATTTCCAGAAATGGTGT 639
QY      265  GTCCAAAGAAAGCTACTATCGATACGTTTGCACGAGAACAAATCATGCCGACCAACAAATGCA 324
DB      640  GTTGAGAGACGAAGTATTATCATACCGAGCTCGCGCGGAAGATGTCAAAAGATTGATAGT 699

```

---

```

QY      325  CAACG 329
DB      700  CGACG 704

RESULT 6
LOCUS B0834782
DEFINITION B0834782 686 bp mRNA linear EST 08-AUG-2002
            Po ad_03D04_5' similar to CAB58493 SEQUENCE 5 FROM PATENT
            WO9424281. unidentified, mRNA sequence.
ACCESSION B0834782
VERSION B0834782.1 GI:22139096
KEYWORDS EST.
SOURCE Psoroptes ovis (sheep scab mite)
ORGANISM Psoroptes ovis
          Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
          Acariformes; Sarcoptiformes; Astigmata; Psoroptida; Sarcoptoidea;
          Psoroptidae; Psoroptes.
REFERENCE 1 (bases 1 to 686)
AUTHORS Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and
          Knox,D.P.
TITLE Expressed sequence tag survey of gene expression in the scab mite
JOURNAL Psoroptes ovis--allergens, proteases and free-radical scavengers
MEDLINE Parasitology 126 (Pt 5), 451-460 (2003)
PUBMED 22677800
COMMENT Contact: Blaxter ML
          Institute of Cell, Animal and Population Biology
          University of Edinburgh
          Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
          3JT, UK.
          Tel: +44 131 650 6760
          Fax: +44 131 670 5450
          Email: mark.blaxter@ed.ac.uk
          The library was prepared for Dr David Knox, Moredun Research
          Institute, UK Sequencing was performed by Mark Welsh, ICAPB,
          Edinburgh
PCR Primers
FORWARD: M13R
BACKWARD: T7PL
Plate: 01 row: D column: 04
High quality sequence stop: 542.
FEATURES             source
    source
        Location/Qualifiers
            1..686
                /organism="Psoroptes ovis"
                /mol_type="mRNA"
                /db_xref="taxon:83912"
                /clone="Po ad_03D04"
                /sex="Mixed"
                /dev_stage="Mixed (predominantly adult)"
                /clone_lib="Psoroptes ovis mixed"
                /note="Vector: Lambda Triplex2; Site.1: SfilA; Site.2:
                SfilB; Psoroptes ovis is the infective agent causing sheep
                scab, a serious dermatitis in sheep. The cDNA was
                obtained from mites harvested from the skin of previously
                unexposed sheep harbouring a moderate challenge
                infection. The cDNA was cloned using the Clontech Lambda
                Triplex2 Smart cDNA cloning system."
ORIGIN
Query Match      20.5%; Score 136.8; DB 5; Length 686;
Best Local Similarity 67.6%; Pred. No. 5.7e-31;
Matches 192; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY      25  AATGCTCCAGCTGAATCGATTTGCGACAATCGCAACTGTCACCTCCCATTCGTATGCAA 84
DB      403  AATTACACACGAAATGATTTACGTGCTATGGGTGATGTCACAAAATCAAGAATCAA 462
QY      85  GGAGGCTGTGGTTCATGTTGGGCTTTCTCTGGTGTGGCGCAACTGCAATCAGCTTATTG 144
DB      463  GCTCATGTGCTCATGCTGGGCTTTCTCCAGTGTGCTACAGTCGANTCAACATCTTA 522

```



```

QY 145 GCTCACCGTAATCAATCATTTGGATCTTCTGAAACAAGAAATTAGTCGATTTGGCTTCCCAA 204
Db 523 GCATACCGTAATGATATAGTATCTTTCTGAACAGAAATTAGTACTGTCATCAACA 582
QY 205 CACGTTGTCATGGTGATACCATTCACCGTGTAFTGAATACATCCAAATGAGTGC 264
Db 583 CACGTTGTCGACGGTGACACAAATGCCACGGTTTGGGTTATATTGCAGAAATGGTGT 642
QY 265 GTCGAAGAAGCTACTATTCGATACGTTCCAGCAGAACAAATCATG 308
Db 643 GTTGAAGACGAAGTATTCATACCGAGTCCGCCGGAAGATG 686

RESULT 7
LOCUS BQ834852
DEFINITION Po ad_04C02_5, 638 bp mRNA linear EST 08-AUG-2002
Po ad_04C02_5, similar to CAB58493 SEQUENCE 5 FROM PARENT
WO9424281.1. unidentified, mRNA sequence.
ACCESSION BQ834852
VERSION BQ834852
KEYWORDS GI:22139166
SOURCE Psoroptes ovis (sheep scab mite)
ORGANISM Psoroptes ovis
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Acariiformes; Sarcoptiformes; Astigmata; Psoroptidia; Sarcopitoidea;
Psoroptidae; Psoroptes.
REFERENCE 1 (bases 1 to 638)
AUTHORS Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and
Knox,D.P.
TITLE Expressed sequence tag survey of gene expression in the scab mite
Psoroptes ovis--allergens, proteases and free-radical scavengers
JOURNAL Parasitology 126 (Pt 5), 451-460 (2003)
MEDLINE 22677800
PUBMED 12793649
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared for Dr David Knox, Moredun Research
Institute, UK Sequencing was performed by Mark Welsh, ICAPB,
Edinburgh
PCR Primers
FORWARD: M13R
BACKWARD: T7PL
Plate: 04 row: C column: 02
High quality sequence stop: 487.
Location/Qualifiers
1..638
/organism="Psoroptes ovis"
/mol_type="mRNA"
/db_xref="taxon:83912"
/clone="Po_ad_04C02"
/sex="Mixed"
/dev_stage="Mixed (predominantly adult)"
/clone_lib="Psoroptes ovis mixed"
/note="Vector: Lambda Triplex2; Site 1: SfIIA; Site 2:
SfiIB; Psoroptes ovis is the infective agent causing sheep
scab, a serious dermatitis in sheep. The cDNA was
obtained from mites harvested from the skin of previously
unexposed sheep harbouring a moderate challenge
infection. The cDNA was cloned using the Clontech Lambda
Triplex2 Smart cDNA cloning system."

FEATURES
source
1..638
/organism="Psoroptes ovis"
/mol_type="mRNA"
/db_xref="taxon:83912"
/clone="Po_ad_04C02"
/sex="Mixed"
/dev_stage="Mixed (predominantly adult)"
/clone_lib="Psoroptes ovis mixed"
/note="Vector: Lambda Triplex2; Site 1: SfIIA; Site 2:
SfiIB; Psoroptes ovis is the infective agent causing sheep
scab, a serious dermatitis in sheep. The cDNA was
obtained from mites harvested from the skin of previously
unexposed sheep harbouring a moderate challenge
infection. The cDNA was cloned using the Clontech Lambda
Triplex2 Smart cDNA cloning system."

ORIGIN
Query Match 16.8%; Score 112; DB 5; Length 638;
Best Local Similarity 71.2%; Pred No. 2, 9e-23;
Matches 148; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

```

QY 25 AATGCTCCAGTCGAAATCGATTTGCGACAAATGCGAACTGTCACCTCCCATTCGATGCAA 84
Db 429 AATTATACACACAGAAATGATTTACGTGCTATGGTCATGTCACAAAATCAAGATCAA 488
QY 85 GGAGGCTGTGGTTCATGTTGGGCTTTCTCTGGTGTTCGCCCACTGAATCAGCTTTATTTG 144
Db 489 GCTGCATGTGGTTCATGCTGGGCTTTCTCCAGTGTGTGTACAGTCGAATCAACATACTTA 548
QY 145 GTCACCGTAAATCAATCATTCGATCTTGTGTAACAGAAATAGTCGATTTGCTTCCCAA 204
Db 549 GCATACCGTAATGATCATTTAGATCTTCTGAAACAAGAAATTAGTACTGTCATCACA 608
QY 205 CACGTTGTCATGGTGATACCATTCAC 232
Db 609 CACGTTGCGACGGTGACACAAATGCCCC 636

RESULT 8
LOCUS BQ835052
DEFINITION Po ad_06F06_5, 612 bp mRNA linear EST 08-AUG-2002
Po ad_06F06_5, similar to CAB58493 SEQUENCE 5 FROM PARENT
WO9424281.1. unidentified, mRNA sequence.
ACCESSION BQ835052
VERSION BQ835052
KEYWORDS GI:22139366
SOURCE Psoroptes ovis (sheep scab mite)
ORGANISM Psoroptes ovis
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Acariiformes; Sarcoptiformes; Astigmata; Psoroptidia; Sarcopitoidea;
Psoroptidae; Psoroptes.
REFERENCE 1 (bases 1 to 612)
AUTHORS Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and
Knox,D.P.
TITLE Expressed sequence tag survey of gene expression in the scab mite
Psoroptes ovis--allergens, proteases and free-radical scavengers
JOURNAL Parasitology 126 (Pt 5), 451-460 (2003)
MEDLINE 22677800
PUBMED 12793649
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The library was prepared for Dr David Knox, Moredun Research
Institute, UK Sequencing was performed by Mark Welsh, ICAPB,
Edinburgh
PCR Primers
FORWARD: M13R
BACKWARD: T7PL
Plate: 06 row: F column: 06
High quality sequence stop: 549.
Location/Qualifiers
1..612
/organism="Psoroptes ovis"
/mol_type="mRNA"
/db_xref="taxon:83912"
/clone="Po_ad_06F06"
/sex="Mixed"
/dev_stage="Mixed (predominantly adult)"
/clone_lib="Psoroptes ovis mixed"
/note="Vector: Lambda Triplex2; Site 1: SfIIA; Site 2:
SfiIB; Psoroptes ovis is the infective agent causing sheep
scab, a serious dermatitis in sheep. The cDNA was
obtained from mites harvested from the skin of previously
unexposed sheep harbouring a moderate challenge
infection. The cDNA was cloned using the Clontech Lambda
Triplex2 Smart cDNA cloning system."

FEATURES
source
1..612
/organism="Psoroptes ovis"
/mol_type="mRNA"
/db_xref="taxon:83912"
/clone="Po_ad_06F06"
/sex="Mixed"
/dev_stage="Mixed (predominantly adult)"
/clone_lib="Psoroptes ovis mixed"
/note="Vector: Lambda Triplex2; Site 1: SfIIA; Site 2:
SfiIB; Psoroptes ovis is the infective agent causing sheep
scab, a serious dermatitis in sheep. The cDNA was
obtained from mites harvested from the skin of previously
unexposed sheep harbouring a moderate challenge
infection. The cDNA was cloned using the Clontech Lambda
Triplex2 Smart cDNA cloning system."

ORIGIN

```

Query Match 16.8%; Score 111.6; DB 5; Length 612;  
 Best Local Similarity 71.4%; Pred. No. 3.8e-23; Indels 0; Gaps 0;  
 Matches 147; Conservative 0; Mismatches 59;

QY 25 AATGCTCCAGCTGAAATCGATTGGCAGCAAAATGCGAACTGTCACTCCCATTCGTATGCAA 84  
 Db 406 AATTTACCACACGAAATGATTACGTGTATGGTCAATGTCACAAAATCAAGAATCAA 465

QY 85 GAGGGCTGTGCTCATCTGTTGGCTTTCTCTGTTGTCGCCGAATCAATCAGCTTATTG 144  
 Db 466 GCTGCATGTGGCTCATCTGTTGGCTTTCTCTGTTGTCGCCGAATCAATCAGCTTATTG 525

QY 145 GCTCACCAGTAATCAATCAATGATGCTGTGTAACAGAAATAGTGTGCTTCCCAA 204  
 Db 526 GCATACCGTAATGATCAATGATGCTTTCTGACAGAAATAGTGTGCTGTCATCAAA 585

QY 205 CACGGTGTGTCATGCTGATACCAATCC 230  
 Db 586 CACGGTGTGTCATGCTGATACCAATCC 611

RESULT 9  
 CA305277 609 bp mRNA linear EST 01-NOV-2002  
 LOCUS ESSU900 s.scabiei cdna library Sarcophyes scabiei cdna clone  
 DEFINITION SAS0725, mRNA sequence.  
 ACCESSION CA305277  
 VERSION CA305277.1 GI:24468331  
 KEYWORDS EST.  
 SOURCE Sarcophyes scabiei  
 ORGANISM Sarcophyes scabiei  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 Acariformes; Sarcophytes; Astigmata; Psoroptidae; Sarcophytoidea;  
 Sarcophytidae; Sarcophytes.  
 REFERENCE 1 (bases 1 to 609)  
 AUTHORS Ljunggren,E.L., Nilsson,D., Naglund,K. and Mattsson,J.G.  
 TITLE Expressed sequence tag analysis of the parasitic mite Sarcophyes scabiei  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Mattsson J.G.  
 Department of Parasitology (SWEPAR)  
 National Veterinary Institute  
 SE-751 89 Uppsala, Sweden  
 Tel: +46 18 674120  
 Fax: +46 18 674120  
 Email: Jens.mattsson@sva.se  
 Similar to gi|627141|pir|A61500 allergen Der f I precursor -  
 house-dust mite [Dermatophagoides farinae]  
 High quality sequence stop: 609.  
 FEATURES Location/Qualifiers  
 source 1..609  
 /organism="Sarcophyes scabiei"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:52283"  
 /clone="SAS0725"  
 /note="The Sarcophyes scabiei mixed life stage library was constructed by Jens G Mattsson. cDNAs were synthesized from poly(A)+ RNA by oligo d(r) priming, size-selected and directionally cloned into the Uni-ZAP lambda vector (Stratagene). The primary library was amplified on XLI-Blue MRF+ cells."

ORIGIN  
 Query Match 15.4%; Score 102.4; DB 6; Length 609;  
 Best Local Similarity 52.8%; Pred. No. 2.8e-20;  
 Matches 245; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

QY 162 ATTGGATCTTGTGGAACAAAGTAATAGTGTGCTTCCCAACAGGTTGTGATGGTGA 221  
 Db 6 ATGACACATATCCGACACAACTTGTGATTTGCTTGACCAAGGATGTGATGGCGA 65

QY 222 TACATTCACGTGGTATTGTAATACATCCACATATGTTGCTCCCAAGAAAGCTACTA 281

Db 66 AAAACCCACACCGGTTTCAATATCTTCTCGAAAAGGTGTAAACAACCTGGCGATAGATA 125  
 QY 282 TCGATAGCTTGCAGAGACAATCATGCGCACCAAAATGACACAGCTTCGGTATCTC 341  
 Db 126 TCCTTATGTTGGAAGAGTTCAACCTTTCAGAGCCTCCGATCGGTCTATACTAAGATTAG 185  
 QY 342 AAACATTTGCAAAATTTACCCACCAATGCAAAATAATTCGTGAAGCTTTGGCTCAAAAC 401  
 Db 186 ATCGTTTGTGGGTCTATCTCTCTGATCCTAAGAGATACAGTACTCCTCTCAATCG 245  
 QY 402 CCACAGCGCTATTGCCGTCTATTATGGCATCAAGATTTAGACGATTCGGTCAATTATGA 461  
 Db 246 GAAAGCCGCATTCACGACAGTGTGAAATCACAATATATGCAATTCGGTCACTACGA 305  
 QY 462 TGCCCGAACATCATTTCAACGCGA--TAATGTTACCAACCAAACTATCACGCTCTCAA 518  
 Db 306 TGGTAAAGTGTAAATTGAGACCGAGGTAGGGAGGAAACCTTTATCGATGCTGTAA 365  
 QY 519 CATGTTGTTTACAGTAACGACCAAGGTGTGATTTATGATCGTACGAAACAGTTGGGA 578  
 Db 366 CATCGTAGGATATGCGCAATATTTTGGCAAGGATGCTTGATCGTTCGCAACAGTTGGG 425

QY 579 TACCAATTTGGGTGATTAATGTTTACGTTATTTTGTGCGCAACA 622  
 Db 426 TACTAGTGGGTGATAAAGGATATTGCTATGTTCTCAATGAACA 469

RESULT 10  
 BO834840 557 bp mRNA linear EST 08-AUG-2002  
 LOCUS BO834840  
 DEFINITION Po ad 04A10.1\_TEXF1 Psoroptes ovis mixed Psoroptes ovis cdna clone  
 Po ad 04A10.5, similar to CAB58493 SEQUENCE 5 FROM PATENT  
 WO9424281.1. unidentified, mRNA sequence.  
 ACCESSION BO834840  
 VERSION BO834840.1 GI:22139154  
 KEYWORDS EST.  
 SOURCE Psoroptes ovis (sheep scab mite)  
 ORGANISM Psoroptes ovis  
 Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 Acariformes; Sarcophytes; Astigmata; Psoroptidae; Sarcophytoidea;  
 Psoroptidae; Psoroptes.  
 REFERENCE 1 (bases 1 to 557)  
 AUTHORS Kenyon,F., Welsh,M., Parkinson,J., Whitton,C., Blaxter,M.L. and Knox,D.P.  
 TITLE Expressed sequence tag survey of gene expression in the scab mite Psoroptes ovis--allergens, proteases and free-radical scavengers  
 JOURNAL Parasitology 126 (Pt 5), 451-460 (2003)  
 MEDLINE 22677800  
 PUBMED 12793649  
 COMMENT Contact: Blaxter ML  
 Institute of Cell, Animal and Population Biology  
 University of Edinburgh  
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JH, UK.  
 Tel: +44 131 650 6760  
 Fax: +44 131 670 5450  
 Email: mark.blaxter@ed.ac.uk  
 The library was prepared for Dr David Knox, Moredun Research Institute, UK Sequencing was performed by Mark Welsh, ICAPB, Edinburgh  
 PCR Primers  
 FORWARD: M13R  
 BACKWARD: T7PL  
 Plate: 04 row: A column: 10  
 High quality sequence stop: 448.  
 FEATURES Location/Qualifiers  
 source 1..557  
 /organism="Psoroptes ovis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:83912"  
 /clone="Po ad 04A10"  
 /sex="Mixed"

/dev\_stage="Mixed (predominantly adult)"  
/clone\_lib="Psooropes ovis mixed"  
/note="Vector: Lambda Triplex2; Site\_1: SfiIA; Site\_2:  
SfiIB; Psooropes ovis is the infective agent causing sheep  
scab, a serious dermatitis in sheep. The cDNA was  
obtained from mites harvested from the skin of previously  
unexposed sheep harbouring a moderate challenge  
infection. The cDNA was cloned using the Clontech Lambda  
Triplex2 Smart cDNA cloning system."

ORIGIN

Query Match 13.1%; Score 87.4; DB 5; Length 557;  
Best Local Similarity 69.8%; Pred. No. 1.3e-15;  
Matches 118; Conservative 0; Mismatches 51; Indels 0; Gaps 0;  
  
QY 25 AATGCTCAGCTGAATGATTTGGACAAATGGCAATGCACTGCCCATTCGTTATGCAA 84  
Db AATTATACACAGCAATGATTTAGCTGCTATGGGTCAATGACACAAAATCAAGATCAA 448  
  
QY 85 GGAGCTGTGGTTCATGTTGGCTTTCTCTGCTGTGGCGCAACTGAATCAGCTATTG 144  
Db GCTGATGTGCTCAGCTGGCTTTCTCCAGTGTGCTACAGTCGATCACTACTTA 508  
  
QY 145 GCTACCGTATCAATCATTTGGATTTGCTGCTGCAACGAATTAGTCGATT 193  
Db GCATACCGTATGATGATTTAGATCTTTCTGAACAAAGAAATTAGTTGACT 557

RESULT 11

BU792027 953 bp mRNA linear EST 23-OCT-2003  
LOCUS SF2AGF02 SJF Schistosoma japonicum cDNA similar to gb|AAA87849.1|  
DEFINITION (U38476) preprocathesin cathepsin L [Schistosoma japonicum], mRNA  
sequence.  
ACCESSION BU792027.1 GI:28349034  
VERSION BU792027  
KEYWORDS EST.  
SOURCE Schistosoma japonicum  
ORGANISM Schistosoma japonicum  
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.  
REFERENCE 1 (bases 1 to 953)  
AUTHORS Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R.,  
Wang, Z.J., Kong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J.,  
Xu, X.N., Wang, S.X., Fu, G., Zhang, X.L., Wang, Z.Q., Brindley, P.J.,  
McManus, D.P., Xue, C.L., Feng, Z., Chen, Z. and Han, Z.G.  
Evolutionary and biomedical implications of a Schistosoma japonicum  
complementary DNA resource  
Nat. Genet. 35 (2), 139-147 (2003).  
JOURNAL 22879925  
MEDLINE 12973349  
PUBMED  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn.

FEATURES

source  
1. .953  
Location/Qualifiers  
/organism="Schistosoma japonicum"  
/mol\_type="mRNA"  
/db\_xref="taxon:6182"  
/sex="female"  
/tissue\_type="Whole body"  
/dev\_stage="adult"  
/lab\_host="rabbits"  
/clone\_lib="SJF"

ORIGIN

Query Match 10.7%; Score 71.4; DB 5; Length 953;  
Best Local Similarity 48.9%; Pred. No. 1.4e-10;  
Matches 192; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 41 TCGATTTGGACAAATGGCAACTGTCTCCATTCGTTATGCAAGAGGCTGTGGTTCAAT 100  
Db TCGATTTGGAGATAATGACACAGTAACACAGTGAAGACACAAATAATTTGGATCTT 452  
  
QY 101 GTTGGGCTTTCTCTGCTGTGTCCGCAACTGAATFCAGCTTATTTGGCTCCCGTAATCAAT 160  
Db CATGGGCATTTGCTGTGTGGAAGCATTAGAGGGCAACTCAATTAAGACAAATAAAC 512  
  
QY 161 CATGGATCTTGCTGAACAAAGAAATTAGTCGATTTGCTTCCCAACACGGTTGTCTGGTG 220  
Db TTATTCATTTATCTGCACAAACATTTGATTTGTTACTGGAGATCATGAATTTGTGGAGA 572  
  
QY 221 ATACCATTCACGTGTTATTAATACATCAACATAATGTTGTCGTCGTCGAAGAAGCTACT 280  
Db ATCCATTACAGTTGATTTGACTACATATAACACAAATGTTGTTGAGTCTGAGGATGATT 632  
  
QY 281 ATGCATACGTTGCACAGAACAAATCATGCGACCAACCAATGCACACGTTTCGGTATCT 340  
Db ATAAGTTCGTTGGAATTTAGAAAATTTACTTACAAATGCATCAAAAAGTCGTCATTACAG 692  
  
QY 341 CAACATTTGCCAAATTTACCCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAA 400  
Db CATCTCATACTCACAAGTTTACCAATTAGTGAAATGAACCTTCAACAGCTTCGTATA 752  
  
QY 401 CCCACAGCGCTATTGCGCTCATTTATTGGCATCA 433  
Db CATATGGAACCTATCGCAGTAACAATTCGAATGA 785

RESULT 12

BG583083 740 bp mRNA linear EST 11-APR-2001  
LOCUS EST494833 GVN Medicago truncatula cDNA clone pgVN-71H21 5' end,  
DEFINITION mRNA sequence.  
ACCESSION BG583083  
VERSION BG583083.1 GI:13598147  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
Medicago.  
REFERENCE 1 (bases 1 to 740)  
AUTHORS Fedorova, M., Pierson, B.L., Samad, D.A., Vance, C.P., Gantt, G.S.,  
Town, C.D., Van Aken, S., Utterback, T., Cho, J. and Fraser, C.M.  
ESTs from one month old nitrogen-fixing root nodules of Medicago  
truncatula, 2001  
Unpublished (2001)  
Contact: Carroll P. Vance  
Department of Agronomy and Plant Genetics  
University of Minnesota  
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA  
Tel: 612 625 5715  
Fax: 651-649-5058  
Email: vance004@maroon.tc.umn.edu  
University of Minnesota name: M384159e TIGR sequence name:  
MTCCE477K More information is available at: http://www.medicago.org  
Seq primer: Skmod (CTA gAA CTA gtg GAT CC).

FEATURES

source  
1. .740  
Location/Qualifiers  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="genotype Al7"  
/db\_xref="taxon:3880"  
/clone="pgVN-71H21"  
/tissue\_type="N2-fixing root nodules"  
/dev\_stage="effective root nodules harvested one month  
post\_inoculation with Sinorhizobium meliloti"  
/lab\_host="E. coli strain XLOIR"  
/clone\_lib="GVN"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:

XhoI; cDNA was prepared from polyA<sup>+</sup> enriched RNA from effective root nodules harvested one month post inoculation with *Sinorhizobium meliloti*. The cDNA was directionally ligated into the Uni-ZAP XR vector from Strategene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XL0LR cells."

## ORIGIN

Query Match 10.7%; Score 71.2; DB 4; Length 740;  
Best Local Similarity 54.2%; Pred. No. 1.5e-10;  
Matches 167; Conservative 0; Mismatches 138; Indels 3; Gaps 1;  
QY 8 CTTGCAATCAATGGAATGCTCCAGCTGAATTCGACAAATGCGAAATGCGAATGTCA 67  
Db 324 CTTCAACITGATGATGATGTTCCACAAATTTGATTGGAGACAAACAGGAGCTGTA 383  
QY 68 CTCCTATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTCTCTGTGTTCGGCAA 127  
Db 384 CTGATGTTAAGACCAAGGAGCTGTGATGTTGTTGGGCATTTTCAGTCGTGGCTGTG 443  
QY 128 CTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTCGATCTTGTGCAACAGAAATTAG 187  
Db 444 TAGAGGTGCTGTGAAATCAACACTGGCGAGTTGATCTCATTAATCTGAGCAACATTTGG 503  
QY 188 TCGATTGTGTTCCCAACAC---GGTTGTCATGGTGATACCATTCACGTGTGATTTGAAT 244  
Db 504 TTGACTGTGATGAACGGAACCTCTGGGTGTCATGGAGGTAATATGAGTAGTGCCTTCAAT 563  
QY 245 ACATCCAAACATATGTTGCTGCTCAAGAAAGCTACTATCATGATCGTTGCAACGAGAACAT 304  
Db 564 ATATATACAAAAGGTATTTGTTAGTGAACGAGATTATCATACCAAGAGTAGCCAGA 623  
QY 305 CATGCCGA 312  
Db 624 CATGTCAA 631

RESULT 13  
CA996057  
LOCUS  
DEFINITION  
r0906.v1 Meloidogyne hapla J2 pAMP1 v1 Meloidogyne hapla cDNA 5', similar to TR:Q21810 Q21810 R07E3.1 PROTEIN. [1]; mRNA sequence.  
ACCESSION  
CA996057  
VERSION  
CA996057.1 GI:27540928  
KEYWORDS  
EST.  
SOURCE  
Meloidogyne hapla  
ORGANISM  
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida; Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.  
REFERENCE  
1 (bases 1 to 661)  
McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marr, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. J2 were provided by Dr. Valerie Williamson of the University of California at Davis (vwilliamson@ucdavis.edu).  
Seq primer: -40RP from Gibco

High quality sequence stop: 412.

FEATURES  
source

1..661  
Location/Qualifiers  
/organism="Meloidogyne hapla"  
/mol\_type="mRNA"  
/db\_xref="taxon:6305"  
/dev\_stage="J2"  
/lab\_host="DH10B"  
/clone\_lib="Meloidogyne hapla J2 pAMP1 v1"  
/note="Vector: pAMP1 (Gibco); Site 1: NotI; Site 2: SalI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dyna). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1. J2 were provided by Dr. Valerie Williamson of the University of California at Davis (vwilliamson@ucdavis.edu)."

## ORIGIN

Query Match 10.6%; Score 70.4; DB 6; Length 661;  
Best Local Similarity 52.7%; Pred. No. 2.6e-10;  
Matches 177; Conservative 0; Mismatches 156; Indels 3; Gaps 1;  
QY 12 CAGTATCAATGGAATGCTCCAGCTGAATTCGACAAATGCGAAATGCGAATGTCACTCC 71  
Db 148 CAATATTAAACAGAACATCCAAGTCATTTTGACTGCGCTGTAAGAGTAGTATACACC 207  
QY 72 CATTCTGATCAAGAGGCTGTGTTTCATGTTGGGCTTCTCTGGTGTCCGCAACTGA 131  
Db 208 TGTTAAAGCAGGAAATGTGATCTTCTGGCATTTGCTTCTGTCGAACGACAGA 287  
QY 132 ATCAGCTTATTTGGTCAACGTAATCAATCATTCGATCTGCTGACAGAAATTAATCGA 191  
Db 268 ATCTGCATATGCTGTAGCATCATGGAGTTTGAAGAGTTTATCAGACAGGAATGTGGA 327  
QY 192 TTGT---GCTTCCACACACGTTGTCTATGTGTATACCATTCACGTTGTTAATATACAT 248  
Db 328 CTGTAACTAGAGATAATCTTGTATGGGGTAATGTGGGAAGGCTTTAGTTTAT 387  
QY 249 CCAACATAATGTTGTCGTCGAAGAAAGCTACTATCATGATGTTGCGACGAGAAATCATG 308  
Db 388 CCAGCATACTGGATTAGTCTCAGAGCGGAATATCTTATGTAGCAGCAGCACTGC 447  
QY 309 CCGACCAACCAAAATGCAACAGCTTTCGTTATCTCAAA 344  
Db 448 CTGTCTGATGAACGACACAGCAAAATTTAACAAA 483

RESULT 14  
BO090548  
LOCUS  
DEFINITION  
r066h04.v1 Meloidogyne hapla egp pAMP1 v1 Meloidogyne hapla cDNA 5', similar to TR:Q26888 Q26888 CATHEPSIN L-LIKE CYSTEINE PROTEINASE.  
ACCESSION  
BO090548  
VERSION  
BO090548.1 GI:20064749  
KEYWORDS  
EST.  
SOURCE  
Meloidogyne hapla  
ORGANISM  
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida; Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.  
REFERENCE  
1 (bases 1 to 420)  
McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marr, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
The library was constructed by Claire Murphy and Dr. James McCarter  
at Washington University, St. Louis. DNA Sequencing by: Washington  
University Genome Sequencing Center  
Seq primer: -40RP from Gibco.  
Location/Qualifiers  
1. .420  
/organism="Meloiodogyne hapla"  
/mol\_type="mRNA"  
/db\_xref="taxon:6305"  
/dev\_stage="parasitic adult females"  
/lab\_host="DH108"  
/clone\_lib="Meloiodogyne hapla egg pAMP1 v1"  
/note="Vector: pAMP1 (Gibco); Site 1: NotI; Site 2: SalI;  
The library was constructed by Claire Murphy and Dr. James  
McCarter at Washington University, St. Louis. The cDNA was  
made by using Dynabead oligo-dT priming (Dynal). PCR based  
library using a modified protocol from the SMART PCR cDNA  
Synthesis Kit from Clontech. Directionally cloned into the  
UDG sites of pAMP1."

FEATURES  
source

ORIGIN  
Query Match 10.5%; Score 70; DB 5; Length 420;  
Best Local Similarity 53.5%; Pred. No. 2.9e-10;  
Matches 170; Conservative 0; Mismatches 145; Indels 3; Gaps 1;  
QY 12 CAGTATCAATGAATGCTCCAGCTGAATGATTTGGGCAAAATGCGAATGTCACATCC 71  
Db 97 CAATATTAAACAGAACATCCAAATGCTTTGACTGGGCTGTGAAGGAGTAGTTACACC 156  
QY 72 CATTCGTATGCAAGGAGGCTGTGTTTCATGTTGGGCTTCTCTGTTGTCGCAATCA 131  
Db 157 TGTAAAGCAGGGAAGAAATGGAATCTGTTGGCAATTTGCTCTGTTGCAACAGACA 216  
QY 132 ATCAGCTATTGTTGCTACCGGTAATCAATCAATGGAATCTGCTGAACAAAGATTAGTCA 191  
Db 217 ATCTGCATATGCTGTAGCACATGGAGTTTGAAGGAGTTTATCAGAACAGGAATGTTGA 276  
QY 192 TTGT---GCTTCCCAACAGGTTGTCATGATACCATCCATCCAGTGTATGATACAT 248  
Db 277 CTGTAACCTAGAGATAATGCTTTGTAATGGGGTAATGTTGAAGGGCTTTTGTATTTAT 336  
QY 249 CCAACATAATGTTGCTGCCAAGAAAGTACTATCGATACGTTGACGAGAACATCATG 308  
Db 337 CCACGATACCTGGATTAGTCTCAGAGGGCAATATCCTTATGTAGCACACCGACACCTGC 396  
QY 309 CCGACCAACCAATGACCA 326  
Db 397 CTGCTGTGATGAACGAACA 414

RESULT 15  
BU716792

LOCUS  
DEFINITION  
BU716792 734 bp mRNA linear EST 23-OCT-2003  
SJM2AXB11 SJM Schistosoma japonicum cDNA similar to gb|AA87849.1|  
(U38476) preprocathepsin cathepsin L [Schistosoma japonicum], mRNA  
sequence.

ACCESSION

BU716792

VERSION

BU716792.1 GI:28324139

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Schistosoma japonicum  
Schistosoma japonicum  
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.  
1 (bases 1 to 734)  
Hu, W., Yan, Q., Shen, D. K., Liu, F., Zhu, Z. D., Song, H. D., Xu, X. R.,  
Wang, Z. J., Rong, Y. P., Zeng, L. C., Wu, J., Zhang, X., Wang, J. J.,  
Xu, X. N., Wang, S. Y., Fu, G., Zhang, X. L., Wang, Z. Q., Brindley, P. J.,

McManus D.P., Xue, C.L., Feng, Z., Chen, Z. and Han, Z.G.  
Evolutionary and biomedical implications of a Schistosoma japonicum  
complementary DNA resource  
Nat. Genet. 35 (2), 139-147 (2003)  
22879925  
PUBMED  
12973349  
COMMENT  
Contact: Zequang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn.  
Location/Qualifiers  
1. .734  
/organism="Schistosoma japonicum"  
/mol\_type="mRNA"  
/db\_xref="taxon:6182"  
/sex="male"  
/tissue\_type="Whole body"  
/dev\_stage="adult"  
/lab\_host="rabbits"  
/clone\_lib="SJM"

FEATURES  
source

ORIGIN

Query Match 10.4%; Score 69.2; DB 5; Length 734;  
Best Local Similarity 51.9%; Pred. No. 6.3e-10;  
Matches 152; Conservative 0; Mismatches 141; Indels 0; Gaps 0;  
QY 16 ATCAATGGAATGCTCCAGCTGAATCGATTTGGGCAAAATGCGAATGTCACCTCCCAT 75  
Db 411 ATGATGAAGAATTAACCGTTACATTTGATGGGCAATTTGCTGTGGAAGCATTAAGAAG 470  
QY 76 CGTATGCAAGGAGGCTGTGTTTCATGTTGGGCTTTCTCTGTTGTTGCCGAATGAAATCA 135  
Db 471 AAGACACAGATAATTTGGATCTTCATGGGCAATTTGCTGTGGAAGCATTAAGAAGG 530  
QY 136 GCTTATTTGCTCACCAGTAATCAATCATTTGGATCTTTGCTGAACAAAGATTAGTCGATT 195  
Db 531 CAATCANATTAAGAACAAATAACTTATTCATTTATCTGCACAAACAATTTGATTGT 590  
QY 196 GCTTCCCAACACCGTTGTCTATGGTGATACCATTCACAGTGGTATTGAATACATCCACAT 255  
Db 591 ACTGGAGATCATGAATGTTGAGATCCATTACAGGTTGGATTGACTACATAAACAC 650  
QY 256 AATGGTGTCTCCCAAGAAAGCTACTATCGATACGTTGCGAGGAAACAATCATG 308  
Db 651 AATGGTGTGAGTCTGAGGATAATTATAGTTCGTTGGAATGTAGAAATTTG 703

Search completed: October 29, 2004, 01:43:31  
Job time : 6948.1 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 19:39:58 ; Search time 825.956 Seconds  
(without alignments)  
4232.808 Million cell updates/sec

Title: US-09-867-159A-1

Perfect score: 666

Sequence: 1 actaacgctgcagatcaaa.....atccatattgttcattcttc 666

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	666	6 AAL41281	AAL41281 DNA encod
2	666	100.0	666	6 ABQ80833	ABQ80833 Cysteine
3	664.4	99.8	834	2 AAG41211	AAG41211 Dermatoph
4	664.4	99.8	834	2 AAG62308	AAG62308 Der p I c
5	664.4	99.8	834	2 AAG79618	AAG79618 Sequence
6	664.4	99.8	834	2 AAV61384	AAV61384 Dermatoph
7	664.4	99.8	834	2 AAZ23906	AAZ23906 Dermatoph
8	664.4	99.8	834	2 AAS30721	AAS30721 House dus
9	664.4	99.8	857	2 AAG23373	AAG23373 cDNA codi
10	664.4	99.8	857	2 AAG58665	AAG58665 Protein a
11	664.4	99.8	857	6 ABQ56342	ABQ56342 cDNA enco
12	664.4	99.8	1172	2 AAG58669	AAG58669 Protein a
13	662.8	99.5	756	12 AAK52165	AAK52165 Partial A
14	662.8	99.5	1089	12 ADK52157	ADK52157 Full leng
15	662.8	99.5	1099	12 ADK52164	ADK52164 Full leng
16	661.2	99.3	666	6 ABK95639	ABK95639 House dus
17	661.2	99.3	669	5 AAH22337	AAH22337 D. pteron
18	661.2	99.3	669	5 AAH22336	AAH22336 D. pteron
19	661.2	99.3	669	12 AAK52139	AAK52139 Der pl al
20	661.2	99.3	849	12 ADL27354	ADL27354 Oligonuc
21	661.2	99.3	906	5 AAH22334	AAH22334 D. pteron

C	22	661.2	99.3	906	5	AAH22335	D. pteron
23	661.2	99.3	909	6	AAI1090	DNA encod	
24	661.2	99.3	950	5	AAH22332	D. pteron	
C	25	661.2	99.3	960	5	AAH22333	D. pteron
26	661.2	99.3	1099	10	ABQ80213	Der p1 co	
27	661.2	99.3	1099	12	ADM57313	Modular a	
28	660.8	99.2	963	12	AAV72472	D. pteron	
29	659.6	99.0	1089	12	ADM57301	Modular a	
30	658.2	98.8	951	2	AAV72474	D. pteron	
31	658	98.8	963	2	AAV72473	D. pteron	
32	656.4	98.6	963	2	AAV72475	D. pteron	
C	33	652.2	97.9	669	5	AAH22376	D. pteron
C	34	652.2	97.9	909	5	AAH22372	D. pteron
35	650	97.6	672	5	AAH22375	D. pteron	
36	650	97.6	912	5	AAH22371	D. pteron	
C	37	649	97.4	666	5	AAH22386	D. pteron
38	649	97.4	666	5	AAH22385	D. pteron	
C	39	649	97.4	906	5	AAH22382	D. pteron
40	649	97.4	906	5	AAH22381	D. pteron	
41	644.2	96.7	906	5	AAH22389	D. pteron	
C	42	644.2	96.7	906	5	AAH22390	D. pteron
43	640.4	96.2	666	6	ABK95644	House dus	
44	640.4	96.2	666	6	ABK95641	House dus	
45	638.8	95.9	666	6	ABK95642	House dus	

ALIGNMENTS

RESULT 1  
AAL41281  
ID AAL41281 standard; DNA; 666 BP.  
XX  
AC AAL41281;

DT 02-JAN-2003 (first entry)

XX DNA encoding a cysteine protease protein.

DE Antiallergic; antiinflammatory; antiasthmatic; dermatologic; allergen;  
XX anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopic eczema; gene; ds.  
XX Dermatophagoides pteronyssinus.

OS Dermatophagoides pteronyssinus.

XX Key Location/Qualifiers

FT CDS 1..666  
FT /\*tag= a  
FT /partial  
FT /product= "Cysteine protease protein"  
FT /transl\_except= (pos:223..225, aa:Arg)  
FT /note= "No start or stop codon"

WO200278736-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-FR001098.

30-MAR-2001; 2001FR-00004370.

03-MAY-2001; 2001FR-00005929.

29-MAY-2001; 2001US-00867159.

(ANTI-) ANTALIS SARL.

Loria E, Terrasse G, Trehin Y;

WPI; 2002-750636/81.

P-PSDB; AAO20568.

Antiallergic compositions containing an anti-histamine, a histamine synthesis inhibitor, and optionally an allergen or nucleic acid coding

PT for the allergen.

PS Claim 18; Page 30; 32pp; French.

XX The invention relates to antiallergic compositions containing an anti-histamine, a histamine synthesis inhibitor, and optionally an allergen or isolated nucleic acid molecule that has at least one polynucleotide sequence coding for the allergen, together with a pharmaceutical carrier. CC The pharmaceutical composition of the invention is useful as a non-specific antiallergic treatment, and also useful in the treatment of allergic hypersensitivity, allergic asthma, allergic rhinitis, and allergic atopic eczema. CC DNA encoding a cysteine protease protein relating to the antiallergic compositions of the invention

XX Sequence 666 BP; 192 A; 150 C; 136 G; 188 T; 0 U; 0 Other;

Query Match 100.0%; Score 666; DB 6; Length 666;

Best Local Similarity 100.0%; Pred. No. 1.5e-199;

Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTAACGCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 60  
DB 1 ACTAACGCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 60  
QY 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 120  
DB 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 120  
QY 121 GCCGCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGCTGAACAA 180  
DB 121 GCCGCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGCTGAACAA 180  
QY 181 GAATTAGTCGATTGTGTTCCCAACAGGTTGTCTGGTGATACCAATTCACGTGATT 240  
DB 181 GAATTAGTCGATTGTGTTCCCAACAGGTTGTCTGGTGATACCAATTCACGTGATT 240  
QY 241 GAATACATCAACATATGTTGCTCCAGAAAGCTTACTGATACGTTCCACGAGAA 300  
DB 241 GAATACATCAACATATGTTGCTCCAGAAAGCTTACTGATACGTTCCACGAGAA 300  
QY 301 CAATCATGCGGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
DB 301 CAATCATGCGGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
QY 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACAGCGCTATTGCCGTC 420  
DB 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACAGCGCTATTGCCGTC 420  
QY 421 ATATTGGCATCAAGATTTAGAGCATTCCTGTCATTATGATGCGCGCAACATCAATCAA 480  
DB 421 ATATTGGCATCAAGATTTAGAGCATTCCTGTCATTATGATGCGCGCAACATCAATCAA 480  
QY 481 CGCGATATGTTTACCAACCAACTATCATCGCTGTCAACATTTTGGTACAGTAACGCA 540  
DB 481 CGCGATATGTTTACCAACCAACTATCATCGCTGTCAACATTTTGGTACAGTAACGCA 540  
QY 541 CAAGGTGTCGATTATTCGATCGTACGAACAGTTGGGATACCAATTTGGGCTGATATGTT 600  
DB 541 CAAGGTGTCGATTATTCGATCGTACGAACAGTTGGGATACCAATTTGGGCTGATATGTT 600  
QY 601 TACGGTATTTTGTGTCACATCGATTGATGATGATTTGAAGAATATCCATATGTTGTC 660  
DB 601 TACGGTATTTTGTGTCACATCGATTGATGATGATTTGAAGAATATCCATATGTTGTC 660  
QY 661 ATTCTC 666  
DB 661 ATTCTC 666

RESULT 2

ABQ80833

JD ABQ80833 standard; cDNA; 666 BP.

XX

AC ABQ80833;

XX 13-DEC-2002 (first entry)

XX Cysteine protease coding sequence.

XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;

XX immunotherapy; allergen; allergic hypersensitivity reaction;

XX allergic asthma; allergic rhinitis; allergic atopic eczema;

XX cysteine protease; gene; ds.

XX Dermatophagoides pteronyssinus.

XX Key Location/Qualifiers

CDS 1..666

FT /\*tag= a

FT /product= "Cysteine protease"

FT /transl\_except= (pos:313..315,aa:Arg)

FT /note= "No start or stop codon given"

XX FR2822709-A1.

XX 04-OCT-2002.

XX 03-MAY-2001; 2001FR-00005929.

XX 30-MAR-2001; 2001FR-00004370.

XX (ANTI-) ANTIALIS SAREL.

XX Loria B, Terrasse G, Trehin Y;

XX WPI; 2002-735037/80.

XX P-PSDB; ABB98533.

XX Antiallergic composition, useful for preventing and treating e.g. asthma,

XX rhinitis or eczema, containing at least two of allergen, antihistamine

XX and histamine synthesis inhibitor.

XX Disclosure; Page 27; 33pp; French.

XX The present invention relates to an antiallergic pharmaceutical

XX composition (I) comprising a pharmaceutical carrier containing an active

XX agent combination of at least two of: an allergen; an antihistamine; and

XX a histamine synthesis inhibitor. (I) is used for treating or preventing

XX allergic hypersensitivity reactions, especially allergic asthma, allergic

XX rhinitis or allergic atopic eczema, in babies, children or adults. The

XX present sequence is the coding sequence for cysteine protease from

XX Dermatophagoides pteronyssinus. The cysteine protease was used as an

XX allergen in the invention

XX Sequence 666 BP; 192 A; 150 C; 136 G; 188 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 666; DB 6; Length 666;

XX Best Local Similarity 100.0%; Pred. No. 1.5e-199;

XX Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTAACGCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 60

DB 1 ACTAACGCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCACAATGCGA 60

QY 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 120

DB 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTCTCTGGTGT 120

QY 121 GCCGCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGCTGAACAA 180

DB 121 GCCGCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGCTGAACAA 180

QY 181 GAATTAGTCGATTGTGTTCCCAACAGGTTGTCTGGTGATACCAATTCACGTGATT 240



```
Db 181 GAATTAGTCGATTGGCTTCCCAACACGTTTGTCTATGGTGATACCAATCCACGTTGTTAT 240
Qy 241 GAATACATCCAAATAATGGTGTCTCCAGAAAGCTACTATCGATAGTTGCGACGAA 300
Db 241 GAATACATCCAAATAATGGTGTCTCCAGAAAGCTACTATCGATAGTTGCGACGAA 300
Qy 301 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
Db 301 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
Qy 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Db 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Qy 421 ATTATTGGCATCAAAAGATTAGACCAATTCGGTCAATTTGATGCGCGAACAATCATTCAA 480
Db 421 ATTATTGGCATCAAAAGATTAGACCAATTCGGTCAATTTGATGCGCGAACAATCATTCAA 480
Qy 481 CCGGATATGTTTACCAACCAACTATCACGCTGTCAACATGTTGTTTACAGTAACGCA 540
Db 481 CCGGATATGTTTACCAACCAACTATCACGCTGTCAACATGTTGTTTACAGTAACGCA 540
Qy 541 CAAGGTGTCGATTATTGGATCGTACGAAACAGTTGGGATACCAATTGGGGTGATAATGTT 600
Db 541 CAAGGTGTCGATTATTGGATCGTACGAAACAGTTGGGATACCAATTGGGGTGATAATGTT 600
Qy 601 TACGTTATTGTTGCGCAACATCGATTGATGATGATGATGATGATGATGATGATGATGATG 660
Db 601 TACGTTATTGTTGCGCAACATCGATTGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 661 ATTCTC 666
Db 661 ATTCTC 666

RESULT 3
AAQ41211
ID AAQ41211 standard; cDNA; 834 BP.
XX
AC AAQ41211;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
DE Dermatophagoides pteronyssinus gene Dep pI.
XX
KW T cell epitope; house dust mite; allergy; soluble; allergen; ss.
XX
OS Dermatophagoides pteronyssinus.
XX
FH Key Location/Qualifiers
FT CDS 1..738
FT /*tag= a
XX
PN WO9308279-A1.
XX
PD 29-APR-1993.
XX
PF 15-OCT-1992; 92MO-US008637.
XX
PR 16-OCT-1991; 91US-00777859.
PR 08-MAY-1992; 92US-00881396.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
WPI; 1993-152472/18.
DR P-PSDB; AAR39359.
XX
PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis
PT and treatment of sensitivity to house dust mite.
XX
```

```
PS Disclosure; Page 45; 176pp; English.
XX
CC The cDNA encoding Dermatophagoides pteronyssinus group 1 allergen Dep pI
CC was obtd. in plasmid form as subclone from lambda gt11 (Chua et al., J.
CC Exp. Med. 167: 175-182 (1988)). The cDNA was used to express a Dep pI
CC allergen which was capable of raising an immune response when contacted
CC with anti-Dep pI antibodies. The protein was used to design a series of
CC overlapping peptides synthesised by standard techniques to cover the
CC whole Dermatophagoides pteronyssinus Dep pI sequence. The T cell epitopes
CC of the protein were mapped by detection of the peptide's ability to
CC stimulate T cell activity. The peptides may be used for diagnosis and
CC treatment of sensitivity to house dust mite allergens. When administered
CC to house dust mite sensitive individuals, the peptides are capable of
CC modifying the allergic response to the allergens. The peptides may be
CC modified for e.g. increasing solubility, enhancing therapeutic or
CC preventive efficacy or stability. See also AAQ41212-6. (Updated on 25-MAR
CC -2003 to correct PN field.)
XX
SQ Sequence 834 BP; 269 A; 170 C; 154 G; 241 T; 0 U; 0 Other;
```

```
Query Match 99.8%; Score 664.4; DB 2; Length 834;
Best Local Similarity 99.8%; Pred. No. 5.2e-199;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 ACTAAGCCCTGCGAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCGACAAATGCCA 60
Db 70 ACTAAGCCCTGCGAGTATCAATGGAATGCTCCAGCTGAAATCGATTGGCGACAAATGCCA 129
Qy 61 ACTGTCACTCCCAATCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTTCTCTGGTGT 120
Db 130 ACTGTCACTCCCAATCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTTCTCTGGTGT 189
Qy 121 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGGTGAACAA 180
Db 190 GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGGTGAACAA 249
Qy 181 GAATTAGTCGATTGTGTTCCCAACACGGTTGTCTGGTGATACCAATTCACGTTGTTAT 240
Db 250 GAATTAGTCGATTGTGTTCCCAACACGGTTGTCTGGTGATACCAATTCACGTTGTTAT 309
Qy 241 GAATACATCCAAATAATGGTGTCTCCAGAAAGCTACTATCGATAGTTGCGACGAA 300
Db 310 GAATACATCCAAATAATGGTGTCTCCAGAAAGCTACTATCGATAGTTGCGACGAA 369
Qy 301 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
Db 370 CAATCATGCCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 429
Qy 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Db 430 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 489
Qy 421 ATTATTGGCATCAAAAGATTAGACGCTTCCGTCATTTATGATGCGCGAACAATCATTCAA 480
Db 490 ATTATTGGCATCAAAAGATTAGACGCTTCCGTCATTTATGATGCGCGAACAATCATTCAA 549
Qy 481 CCGGATAATGTTTACCAACCAACTATCACGCTGTCAACATTTGTTTACAGTAACGCA 540
Db 550 CCGGATAATGTTTACCAACCAACTATCACGCTGTCAACATTTGTTTACAGTAACGCA 609
Qy 541 CAAGGTGTCGATTATTGGATCGTACGAAACAGTTGGATACCAATTTGGGGTGATAATGTT 600
Db 610 CAAGGTGTCGATTATTGGATCGTACGAAACAGTTGGATACCAATTTGGGGTGATAATGTT 669
Qy 601 TACGTTATTGTTGCTGCCAATCGATTTCGATGATGATGATGATGATGATGATGATGATGATG 660
Db 670 TACGTTATTGTTGCTGCCAATCGATTTCGATGATGATGATGATGATGATGATGATGATG 729
Qy 661 ATTCTC 666
Db 730 ATTCTC 735
```



XX W09427634-Al.  
XX 08-DEC-1994.  
XX 01-JUN-1994; 94WO-AU000292.  
XX 02-JUN-1993; 93US-00072832.  
XX (CHIL-) INST CHILD HEALTH RES.  
XX Thomas WR;  
XX WPI; 1995-022467/03.  
XX P-PSDB; AAR66345.  
XX Inducing immunological tolerance with cryptic peptide - esp from allergen  
XX or auto:antigen, esp admin orally.  
XX Disclosure; Page 20-22; 38pp; English.  
XX Immunological tolerance to a protein antigen may be induced by admin. of  
XX a compsn. contg. at least one cryptic peptide derived from that antigen.  
XX BIO mice were immunised with various Der p I peptides and then response  
XX to Der p I and the peptides was measured in the presence of spleen  
XX adherent cells. Peptides contg. amino acids 120-143 and 144-169 of Der p  
XX I were positive i.e., they contained cryptic epitopes. The cDNA fragment  
XX encoding amino acids 131-187 of Der p I was cloned into pGEX and  
XX expressed in bacteria as a fusion protein with glutathione-S-transferase.  
XX Mice were given orally 3 mg of this protein on 3 consecutive days, then 7  
XX days later immunised subcutaneously with native Der p I in complete  
XX Freund's adjuvant. Seven days later lymph nodes, stimulated with protein  
XX or synthetic peptides, were assayed for lymphokines (IL-2). Mice given  
XX the cryptic epitope had a much weaker response than those treated only  
XX with buffer. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 834 BP; 269 A; 170 C; 154 G; 241 T; 0 U; 0 Other;

Query Match 99.8%; Score 664.4; DB 2; Length 834;  
Best Local Similarity 99.8%; Pred. No. 5.2e-199;  
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAAGCCTGAGTATCAATGGAATGCTCAGCTGAATCGATTGCGACAAATGCGA 60  
DB 70 ACTAAGCCTGAGTATCAATGGAATGCTCAGCTGAATCGATTGCGACAAATGCGA 129  
QY 61 ACTGTCACTCCATTCGATGCAAGAGAGCTGTGTTCAATGTTGGGCTTTCTCGTGT 120  
DB 130 ACTGTCACTCCATTCGATGCAAGAGAGCTGTGTTCAATGTTGGGCTTTCTCGTGT 189  
QY 121 GCGGCACTGAATCAGCTTATTGGCTCACCGTAAATCAATCAATCGATTCGTCACAA 180  
DB 190 GCGGCACTGAATCAGCTTATTGGCTCACCGTAAATCAATCAATCGATTCGTCACAA 249  
QY 181 GAATTAGTCGATGTCCTCCCAACACCGGTTGTGATGATACCATTCACGTCGTATT 240  
DB 250 GAATTAGTCGATGTCCTCCCAACACCGGTTGTGATGATACCATTCACGTCGTATT 309  
QY 241 GAATACATCAACATTAATGGTGTGTCAGAAAGCTACTATCGATACGTTTCACGAGAA 300  
DB 310 GAATACATCAACATTAATGGTGTGTCAGAAAGCTACTATCGATACGTTTCACGAGAA 369  
QY 301 CAATCATGCGACACCAATGACCAACGTTTCGGTATCTCAACATTAATGCGCAAAATTAC 360  
DB 370 CAATCATGCGACACCAATGACCAACGTTTCGGTATCTCAACATTAATGCGCAAAATTAC 429  
QY 361 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTTCAACACCAAGCGCTATTGCCGTC 420  
DB 430 CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTTCAACACCAAGCGCTATTGCCGTC 489  
QY 421 ATTATTGGATCAAGATTTAGACGATTCCTCGTATTGATGGCGGACATCATCTTCAA 480  
DB 490 ATTATTGGATCAAGATTTAGACGATTCCTCGTATTGATGGCGGACATCATCTTCAA 549

QY 481 CGCGATAATGTTACCAACCAACTATCAGCTGTCAACATTCGTTGTTACAGTAACGCA 540  
DB 550 CGCGATAATGTTACCAACCAACTATCAGCTGTCAACATTCGTTGTTACAGTAACGCA 609  
QY 541 CAAGGTGTCGATTTATTGGATCGTACGAAACAGTTGGATACCAATTCGGGTGATAATGGT 600  
DB 610 CAAGGTGTCGATTTATTGGATCGTACGAAACAGTTGGATACCAATTCGGGTGATAATGGT 669  
QY 601 TACGGTTATTGCTGCCAACATCGATTTCATGATGATTTGAAGAATATCCATATGTTGTC 660  
DB 670 TACGGTTATTGCTGCCAACATCGATTTCATGATGATTTGAAGAATATCCATATGTTGTC 729  
QY 661 ATTCTC 666  
DB 730 ATTCTC 735  
RESULT 6  
AAV61384  
ID AAV61384 standard; cDNA; 834 BP.  
XX AAV61384;  
AC AAV61384;  
DT 27-AUG-2003 (revised)  
DT 15-MAR-2003 (revised)  
DT 16-DEC-1998 (first entry)  
XX  
DE Dermatophagoides Der p I encoding cDNA.  
XX  
KW genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
KW Der p II; Der p I; Der p II; house dust mite allergy; ss.  
XX  
OS Dermatophagoides.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..738  
FT /\*tag= a  
FT /product= "Der p I"  
XX  
PN USS920862-A.  
XX  
PD 13-OCT-1998.  
XX  
XX 07-JUN-1995; 95US-00482142.  
XX  
PR 14-APR-1993; 93WO-US003471.  
PR 14-APR-1994; 94US-00227772.  
PR 19-MAY-1995; 95US-00445307.  
XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX  
PI Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
PI Shaked Z, Rogers BL;  
XX  
DR WPI; 1998-567590/48.  
DR P-PSDB; AAW71908.  
XX  
XX Dermatophagoides allergen peptides - useful for treating house dust mite  
PT allergy.  
XX  
PS Disclosure; Col 69-72; 155pp; English.  
XX  
CC The present invention describes peptides for treating sensitivity to  
CC house dust mite allergens from the genus Dermatophagoides. Peptides  
CC within the scope of the invention comprise at least one T cell epitope,  
CC or preferably at least two T cell epitopes of a protein allergen selected  
CC from the allergens Der p I, Der p II, Der p I, or Der p II. The invention  
CC also describes modified peptides having similar or enhanced therapeutic  
CC properties as the corresponding, naturally occurring allergen, but having  
CC reduced side effects. The present sequence encodes Der p I from the  
CC present invention. (Updated on 25-MAR-2003 to correct PR field.) (Updated  
CC on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 834 BP; 269 A; 170 C; 154 G; 241 T; 0 U; 0 Other;  
 Query Match 99.8%; Score 564.4; DB 2; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 5.2e-199;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACTAACGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCAAAATGCGA 60  
 Db 70 ACTAACGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCAAAATGCGA 129

Qy 61 ACTGTCTACTCCATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 120  
 Db 130 ACTGTCTACTCCATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 189

Qy 121 GCGCAACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGGATCTTGTGTAACAA 180  
 Db 190 GCGCAACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGGATCTTGTGTAACAA 249

Qy 181 GAATTAGTCGATTTGCTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGGTATT 240  
 Db 250 GAATTAGTCGATTTGCTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGGTATT 309

Qy 241 GAATACATCCAAACATATGTTGCTCCCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 300  
 Db 310 GAATACATCCAAACATATGTTGCTCCCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 369

Qy 301 CAATCATGCGGACCAACCAATGCAACAGCTTTCCGGTATCTCAAACTATTGCCAAATTTAC 360  
 Db 370 CAATCATGCGGACCAACCAATGCAACAGCTTTCCGGTATCTCAAACTATTGCCAAATTTAC 429

Qy 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACAGCGCTATTGCGGTC 420  
 Db 430 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACAGCGCTATTGCGGTC 489

Qy 421 ATTATTTGGCATCAAGATTTAGAGCGATTCGCTCATTTATGATGCGCGAACAATCATTTCAA 480  
 Db 490 ATTATTTGGCATCAAGATTTAGAGCGATTCGCTCATTTATGATGCGCGAACAATCATTTCAA 549

Qy 481 CGCATATGTTTACCAACCAATCATGCGCTGTGCAACATTTGGTGATACGTAACGGA 540  
 Db 550 CGCATATGTTTACCAACCAATCATGCGCTGTGCAACATTTGGTGATACGTAACGGA 609

Qy 541 CAAGTGTCGATTTATGCGATCGTACGAAACAGTTGGGATACCAATTTGGGTGATATGTT 600  
 Db 610 CAAGTGTCGATTTATGCGATCGTACGAAACAGTTGGGATACCAATTTGGGTGATATGTT 669

Qy 601 TACGGTTATTTGTCGCAACATCGATTTGATGATGATTTGAAGAAATATCCATATGTTGTC 660  
 Db 670 TACGGTTATTTGTCGCAACATCGATTTGATGATGATTTGAAGAAATATCCATATGTTGTC 729

Qy 661 ATTCTC 666  
 Db 730 ATTCTC 735

RESULT 7  
 AAZ23906  
 ID AAZ23906 standard; cDNA; 834 BP.  
 XX AC AAZ23906;  
 XX 25-JAN-2000 (first entry)  
 DT Dermatophagoides sp. allergen Der p I cDNA fragment.  
 XX DE  
 XX KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 XX Der p I; ss.  
 XX OS Dermatophagoides sp.  
 XX FH Key Location/Qualifiers

FT CDS  
 FT 1..738  
 FT /\*tag= a  
 FT /product= "Der p I"  
 FT /note= "partial sequence, no start codon given"  
 PN US5968526-A.  
 PD 19-OCT-1999.  
 PF 07-JUN-1995; 95US-00478572.  
 PR 14-APR-1994; 94US-00227772.  
 PR 12-APR-1995; 95WO-US004481.  
 PR 19-MAY-1995; 95US-00445307.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;  
 PI Chen X, Evans S, Kuo M;  
 DR WPI; 1999-590385/50.  
 DR P-PSDB; AAY50356.  
 XX Screening individuals for allergic reactions to T cell epitopes of major  
 PT allergens from house dust mites.  
 XX Disclosure; Col 71-74; 158pp; English.  
 XX This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
 CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the sensitivity  
 CC of the patient to house dust mites. (I) may be used to screen individuals  
 CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
 CC mite is a major cause of a variety of allergic disorders such as asthma,  
 CC rhinitis and ectopic dermatitis. This sequence encodes the house dust  
 CC mite allergen Der p I  
 XX SQ Sequence 834 BP; 269 A; 170 C; 154 G; 241 T; 0 U; 0 Other;

Query Match 99.8%; Score 564.4; DB 2; Length 834;  
 Best Local Similarity 99.8%; Pred. No. 5.2e-199;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACTAACGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCAAAATGCGA 60  
 Db 70 ACTAACGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCAAAATGCGA 129

Qy 61 ACTGTCTACTCCATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 120  
 Db 130 ACTGTCTACTCCATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 189

Qy 121 GCGCAACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGGATCTTGTGTAACAA 180  
 Db 190 GCGCAACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCATTTGGATCTTGTGTAACAA 249

Qy 181 GAATTAGTCGATTTGCTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGGTATT 240  
 Db 250 GAATTAGTCGATTTGCTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGGTATT 309

Qy 241 GAATACATCCAAACATATGTTGCTCCCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 300  
 Db 310 GAATACATCCAAACATATGTTGCTCCCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGGTGT 369

Qy 301 CAATCATGCGGACCAACCAATGCAACAGCTTTCCGGTATCTCAAACTATTGCCAAATTTAC 360  
 Db 370 CAATCATGCGGACCAACCAATGCAACAGCTTTCCGGTATCTCAAACTATTGCCAAATTTAC 429



DT 04-AUG-1992 (first entry)  
 XX cDNA coding for Der p 1, clone p1(13T).  
 DE House dust mite; allergen; desensitisation therapy; sensitivity; ss.  
 KW Dermatophagoides pteronyssinus.  
 XX  
 OS  
 XX Key Location/Qualifiers  
 PH CDS 1..816 /\*tag= a  
 FT polyA\_signal 803..808  
 FT /\*tag= b  
 FT  
 XX WO9204445-A.  
 PN 19-MAR-1992.  
 PD 10-SEP-1991; 91WO-AU000417.  
 PF 11-SEP-1990; 90US-00580655.  
 PR (WAUR-) WEST AUST RES INST.  
 XX Thomas WR, Chua KY;  
 PI WPI; 1992-114353/14.  
 XX P-PSDB; AAR22433.  
 DR  
 XX  
 XX DNA encoding allergens of house dust mite Dermatophagoides - and isolated  
 PT protein allergens and peptide(s) useful in treating and detecting  
 PT sensitivity to mites, esp. D. farinae.  
 XX  
 PS Claim 5; Fig 7; 53pp; English.  
 CC A Dermatophagoides cDNA library was constructed in E. coli Y1090 from  
 CC mRNA isolated from live mites. The library was screened with two probes  
 CC comprising the Der p 1 cDNA BamHI fragments 1-348 and 349- 857. Clones  
 CC expressing the mite allergen Der p 1 were isolated. The Der p 1 cDNA  
 CC insert was amplified from phage DNA by PCR. Each Der p 1 cDNA fragment  
 CC was cloned separately into the M13- derived sequencing vectors mp18 and  
 CC mp19 and sequenced. The protein allergen encoded by the DNA is used as a  
 CC diagnostic reagent in detecting and treating sensitivity to house dust  
 CC mites and in desensitisation therapy. See also AAQ23371.2. (Updated on 27  
 CC -AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 857 BP; 293 A; 170 C; 154 G; 240 T; 0 U; 0 Other;  
 Query Match 99.8%; Score 664.4; DB 2; Length 857;  
 Best Local Similarity 99.8%; Pred. No. 5.3e-199;  
 Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ACTAAGCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60  
 DB 70 ACTAAGCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 129  
 QY 61 ACTGTCACTCCCATTCGATATGCAAGGAGGCTGTGTTTCATGTTGGCTTTCTCTGGTGT 120  
 DB 130 ACTGTCACTCCCATTCGATATGCAAGGAGGCTGTGTTTCATGTTGGCTTTCTCTGGTGT 189  
 QY 121 GCCGCACTCAATCAGCTTTATTTGGCTCACCCTGAAATCAATCATTCGATTCGTAACAA 180  
 DB 190 GCCGCACTCAATCAGCTTTATTTGGCTCACCCTGAAATCAATCATTCGATTCGTAACAA 249  
 QY 181 GAATTAGTCAGTGTGTTCCCAACAGGTTGTTCATGTTGATACCATTCACGTGTATT 240  
 DB 250 GAATTAGTCAGTGTGTTCCCAACAGGTTGTTCATGTTGATACCATTCACGTGTATT 309  
 QY 241 GAATACATCAACATATGTTGTCGTCACCAAGGCTACTTCGATACGTTCCAGGAA 300  
 DB 310 GAATACATCAACATATGTTGTCGTCACCAAGGCTACTTCGATACGTTCCAGGAA 369  
 QY 301 CAATCATCGCGACCAACCAAAATGCAACAGTTTCGGTATCTCAAACTATTTCGCAAAATTAC 360

Db 370 CAATCATCGCGACCAACAAATTCGACAAAGTTTCGGTATCTCAAACTATTGCCAAATTAC 429  
 QY 361 CCACCAATGCAAAACAAATTCGTAAGCTTTGGCTCAAAACCCACAGCGCTATTGCGGTC 420  
 Db 430 CCACCAATGCAAAACAAATTCGTAAGCTTTGGCTCAAAACCCACAGCGCTATTGCGGTC 489  
 QY 421 ATTATTGGCATCAAGATTTAGACGATTCGTCATTATGATGCGCCGAAACAATCATTTCAA 480  
 Db 490 ATTATTGGCATCAAGATTTAGACGATTCGTCATTATGATGCGCCGAAACAATCATTTCAA 549  
 QY 481 CGCGATTAATGTTTACCAACCAAACTATCAAGCTGTCAACATTTGTTGGTTACAGTAACGCA 540  
 Db 550 CGCGATTAATGTTTACCAACCAAACTATCAAGCTGTCAACATTTGTTGGTTACAGTAACGCA 609  
 QY 541 CAAGGTGTGATTTATGATGCTGACGAAACAGTTGGGATACCAATTTGGGTGATATGTT 600  
 Db 610 CAAGGTGTGATTTATGATGCTGACGAAACAGTTGGGATACCAATTTGGGTGATATGTT 669  
 QY 601 TACGTTATTTGCTGCCAATCGATTGATGATGATTGAAGATATCCATATGTTGTC 660  
 Db 670 TACGTTATTTGCTGCCAATCGATTGATGATGATTGAAGATATCCATATGTTGTC 729  
 QY 661 ATTCTC 666  
 Db 730 ATTCTC 735  
 RESULT 10  
 AAQ58665  
 ID AAQ58665 standard; cDNA; 857 BP.  
 XX AC AAQ58665;  
 XX DT 25-MAR-2003 (revised)  
 DT 17-OCT-1994 (first entry)  
 XX Protein allergen of Derp 1.  
 XX KW Der p1; House Dust Mite Allergen; ss.  
 XX OS Dermatophagoides pteronyssinus.  
 XX FH Key Location/Qualifiers  
 FT sig\_peptide 1..69 /\*tag= a  
 FT /product= "partial signal peptide"  
 FT /note= "69 base long 5'proximal end sequence "  
 FT mat\_peptide 70..738 /\*tag= b  
 FT /product= "Der p1 mature protein"  
 FT /note= "encoded in ggt11 p1"  
 FT polyA\_signal 804..809 /\*tag= c  
 XX WO9405790-A1.  
 XX PD 17-MAR-1994.  
 XX PF 10-SEP-1993; 93WO-US008518.  
 PR 10-SEP-1992; 92US-00945288.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Thomas WR, Chua K;  
 XX WPI; 1994-101195/12.  
 DR P-PSDB; AAR47063.  
 XX New protein allergens of house dust mite - used for diagnosing and  
 PT treating sensitivity in an individual to house dust mite allergens.  
 XX

XX  
DT 10-JAN-2003 (first entry)  
XX  
DE cDNA encoding European house dust mite Der p I protein.

```

QY 61 ACTGCTACTCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTTCCTGCTGTT 120
DB 130 ACTGCTACTCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTTCCTGCTGTT 189
QY 121 GCAGCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCAATTCATTCGTCGAACA 180
DB 190 GCAGCACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCAATTCATTCGTCGAACA 249
QY 181 GAATTAGTCGATTTGCTTCCCAACAGGTTGTCTGGTGATACCAATCCAGCTGATTT 240
DB 250 GAATTAGTCGATTTGCTTCCCAACAGGTTGTCTGGTGATACCAATCCAGCTGATTT 309
QY 241 GAATACATCAACATTAATGCTGCTCCCAAGAGGCTACTATCGATACGTTGCAAGAA 300
DB 310 GAATACATCAACATTAATGCTGCTCCCAAGAGGCTACTATCGATACGTTGCAAGAA 369
QY 301 CAATCATGCGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
DB 370 CAATCATGCGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 429
QY 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCAAGAGGCTATTGCGCTC 420
DB 430 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCAAGAGGCTATTGCGCTC 489
QY 421 ATTATTGGCATCAAGATTTAGACGATTCGCTCATTTATGATGGCGCAACATTCATCAA 480
DB 490 ATTATTGGCATCAAGATTTAGACGATTCGCTCATTTATGATGGCGCAACATTCATCAA 549
QY 481 CGCGATAATGGTTACCAACCAACTATACGCTGTCAACATTTGTTGGTTACAGTAACGCA 540
DB 550 CGCGATAATGGTTACCAACCAACTATACGCTGTCAACATTTGTTGGTTACAGTAACGCA 609
QY 541 CAAGGTGTCATATTGATGCTGACGAACAGTTGGGATACCAATTTGGGCTGATATGGT 600
DB 610 CAAGGTGTCATATTGATGCTGACGAACAGTTGGGATACCAATTTGGGCTGATATGGT 669
QY 601 TACGGTTATTTGCTGCCAATCGATTTGATGATGATTTGAAGAAATATCCATATGTTGC 660
DB 670 TACGGTTATTTGCTGCCAATCGATTTGATGATGATTTGAAGAAATATCCATATGTTGC 729
QY 661 ATTCTC 666
DB 730 ATTCTC 735

```

## RESULT 12

AAQ58669 standard; cDNA; 1172 BP.

XX AAQ58669;

XX 25-MAR-2003 (revised)

XX 17-OCT-1994 (first entry)

XX Protein allergen of Der p1.

XX Der p1; House Dust Mite Allergen; ss.

XX Dermatophagoides pteronyssinus.

XX Key Location/Qualifiers

XX sig\_peptide 1..344 a

XX /product= "partial signal peptide"

XX /note= "344 base long 5'proximal end sequence"

XX mat\_peptide 345..1009

XX /tag= b

XX /product= "Der p1 mature protein"

XX /note= "encoded in sg11 p1"

XX polyA\_signal 1078..1083

XX /tag= c

```

PN WO9405790-A1.
XX 17-MAR-1994.
XX 10-SEP-1993; 93WO-US008518.
XX 10-SEP-1992; 92US-00945288.
XX (IMMU-) IMMULOGIC PHARM CORP.
XX Thomas WR, Chua K;
XX WPI; 1994-101195/12.
XX P-PSDB; AAR49920.
XX
XX New protein allergens of house dust mite - used for diagnosing and
XX treating sensitivity in an individual to house dust mite allergens.
XX Example 1; Fig 21; 98pp; English.
XX
XX AAQ58669 is a cDNA clone encoding a full length preproenzyme Der p1. This
XX is encoded in a single open reading frame, with a termination site at
XX position 1011-1013. The mature protein can be used to detect sensitivity
XX in an individual to house dust mite and to reduce the sensitivity of the
XX individual. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1172 BP; 417 A; 227 C; 197 G; 331 T; 0 U; 0 Other;
XX
Query Match 99.8%; Score 564.4; DB 2; Length 1172;
Best Local Similarity 99.8%; Pred. No. 6.2e-199;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACTAACGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATCGCA 60
DB 344 ACTAACGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATCGCA 403
QY 61 ACTGCTACTCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTTCCTGCTGTT 120
DB 404 ACTGCTACTCCCATTCGTATGCAAGAGGCTGTGGTTCATGTTGGGCTTTCCTGCTGTT 463
QY 121 GCCCAACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCAATTCATTCGTCGAACA 180
DB 464 GCCCAACTGAATCAGCTTATTTGGCTCAGCGTAATCAATCAATTCATTCGTCGAACA 523
QY 181 GAATTAGTCGATTTGCTTCCCAACAGGTTGTCTGGTGATACCAATTCACGTGATTT 240
DB 524 GAATTAGTCGATTTGCTTCCCAACAGGTTGTCTGGTGATACCAATTCACGTGATTT 593
QY 241 GAATACATCAACATTAATGCTGCTCCCAAGAGGCTACTATCGATACGTTGCAAGAA 300
DB 584 GAATACATCAACATTAATGCTGCTCCCAAGAGGCTACTATCGATACGTTGCAAGAA 643
QY 301 CAATCATGCGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
DB 644 CAATCATGCGACCAACCAATGCAACAGCTTTCGGTATCTCAAACTATTGCCAAATTTAC 703
QY 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCAAGAGGCTATTGCGCTC 420
DB 704 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCAAGAGGCTATTGCGCTC 763
QY 421 ATTATTGGCATCAAGATTTAGACGATTCGCTCATTTATGATGGCGCAACATTCATCAA 480
DB 764 ATTATTGGCATCAAGATTTAGACGATTCGCTCATTTATGATGGCGCAACATTCATCAA 823
QY 481 CGCGATAATGGTTACCAACCAACTATACGCTGTCAACATTTGTTGGTTACAGTAACGCA 540
DB 824 CGCGATAATGGTTACCAACCAACTATACGCTGTCAACATTTGTTGGTTACAGTAACGCA 883
QY 541 CAAGGTGTCGATTTGATGCTGACGAACAGTTGGGATACCAATTTGGGCTGATATGGT 600
DB 884 CAAGGTGTCGATTTGATGCTGACGAACAGTTGGGATACCAATTTGGGCTGATATGGT 943
QY 601 TACGGTTATTTGCTGCCAATCGATTTGATGATGATTTGAAGAAATATCCATATGTTGC 660

```





XX Recombinant production of acarid protein allergen, useful for diagnosis  
PT and treatment of allergy to house dust mites, comprises growing  
PT transformed eukaryotes, particularly plants.  
XX  
XX Claim 8; SEQ ID NO 19; 55pp; French.  
XX  
XX The present invention relates to a method for production of a recombinant  
CC protein allergen from an acarid of the genera Dermatophagoides or  
CC Euroglyphus. The allergens and also antibodies raised against them, are  
CC useful for diagnosis and treatment of allergies to house dust mites. When  
CC expressed in plants, allergens are synthesized and matured to  
CC biologically active form, with essentially the same pattern of  
CC glycosylation as the native protein. Recombinant expression provides a  
CC pure protein; contrast complex mixtures of allergens currently used. The  
CC present sequence represents the full length Der p1 allergen encoding  
CC sequence.  
XX  
XX Sequence 1089 BP; 338 A; 224 C; 197 G; 330 T; 0 U; 0 Other;  
SQ  
Query Match 99.5%; Score 662.8; DB 12; Length 1089;  
Best Local Similarity 99.7%; Pred. No. 1.9e-198;  
Matches 664; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ACTAACGCTCGAGTATCAATGGAAATGCTCCAGCTGAATCGAATCGAATAATGCGA 60  
Db 343 ACTAACGCTCGAGTATCAATGGAAATGCTCCAGCTGAATCGAATCGAATAATGCGA 402  
QY 61 ACTGTCTACTCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGGCTTCTCTGGTGT 120  
Db 403 ACTGTCTACTCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGGCTTCTCTGGTGT 462  
QY 121 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATGATGATTCGTAACAA 180  
Db 463 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATGATGATTCGTAACAA 522  
QY 181 GAATTAGTCGATTTGGTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGTATT 240  
Db 523 GAATTAGTCGATTTGGTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGTATT 582  
QY 241 GAATACATCAACATAATGTGTGTCGTCACAAAGACTACTATGATACGTTGACGAGAA 300  
Db 583 GAATACATCAACATAATGTGTGTCGTCACAAAGACTACTATGATACGTTGACGAGAA 642  
QY 301 CAATCATGCCGACCAACAAATGCACACGTTTCGGTATCTCAAACTATTGCCAATTTAC 360  
Db 643 CAATCATGCCGACCAACAAATGCACACGTTTCGGTATCTCAAACTATTGCCAATTTAC 702  
QY 361 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
Db 703 CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 762  
QY 421 ATTATGGCATCAAGATTTAGACGATTCGGTATTCATGATGATGCGCCGACATCAATCAA 480  
Db 763 ATTATGGCATCAAGATTTAGACGATTCGGTATTCATGATGATGCGCCGACATCAATCAA 822  
QY 481 CGCGATAATGGTTACCAACCAACTATACCGTGTCAACATTTGTTGGTTACAGTAACGCA 540  
Db 823 CGCGATAATGGTTACCAACCAACTATACCGTGTCAACATTTGTTGGTTACAGTAACGCA 582  
QY 541 CAAGGTGTCATTTATGGATCGTACGAACAGTTGGATACCAATTTGGGTTGATATGGT 600  
Db 883 CAAGGTGTCATTTATGGATCGTACGAACAGTTGGATACCAATTTGGGTTGATATGGT 942  
QY 601 TAGCGTTATTTGCTGCAACATCGATTTGATGATGATTTGAAGAATATCCATATGTTGC 660  
Db 943 TAGCGTTATTTGCTGCAACATCGATTTGATGATGATTTGAAGAATATCCATATGTTGC 1002  
QY 661 ATTCTC 666  
Db 1003 ATTCTC 1008

## RESULT 15

ADK52164  
ID ADK52164 standard; DNA; 1099 BP.  
XX  
AC ADK52164;  
XX  
XX 06-MAY-2004 (first entry)  
DT  
XX Full length ADNC encoding sequence.  
DE  
XX Recombinant protein allergen; Antiallergic; Desensitization; antibody;  
KW allergy; house dust mite; allergen; signal peptide; ds.  
KW  
XX Unidentified.  
OS  
XX WO2004005334-A2.  
PN  
XX 15-JAN-2004.  
PD  
XX  
XX 04-JUL-2003; 2003WO-FR002085.  
PF  
XX 05-JUL-2002; 2002FR-00008485.  
PR  
XX (STAL-) STALLERGENES SA.  
PA (SETE) SOC NAT EXPL IND TABACS & ALLUMETTES.  
XX  
XX Gomord V, Lienard D, Van Ree R, Van Oort E, Dorihaac De Borne F;  
PI Didier Laurent A, Paye L;  
XX  
XX WPI; 2004-083498/08.  
D3  
XX Recombinant production of acarid protein allergen, useful for diagnosis  
PT and treatment of allergy to house dust mites, comprises growing  
PT transformed eukaryotes, particularly plants.  
XX  
XX Disclosure; SEQ ID NO 26; 55pp; French.  
PS  
XX  
XX The present invention relates to a method for production of a recombinant  
CC protein allergen from an acarid of the genera Dermatophagoides or  
CC Euroglyphus. The allergens and also antibodies raised against them, are  
CC useful for diagnosis and treatment of allergies to house dust mites. When  
CC expressed in plants, allergens are synthesized and matured to  
CC biologically active form, with essentially the same pattern of  
CC glycosylation as the native protein. Recombinant expression provides a  
CC pure protein; contrast complex mixtures of allergens currently used. The  
CC present sequence represents full length ADNC encoding sequence.  
XX  
SQ Sequence 1099 BP; 344 A; 225 C; 198 G; 332 T; 0 U; 0 Other;  
Query Match 99.5%; Score 662.8; DB 12; Length 1099;  
Best Local Similarity 99.7%; Pred. No. 1.9e-198;  
Matches 664; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ACTAACGCTCGAGTATCAATGGAAATGCTCCAGCTGAATCGAATCGAATAATGCGA 60  
Db 344 ACTAACGCTCGAGTATCAATGGAAATGCTCCAGCTGAATCGAATCGAATAATGCGA 403  
QY 61 ACTGTCTACTCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGGCTTCTCTGGTGT 120  
Db 404 ACTGTCTACTCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGGCTTCTCTGGTGT 463  
QY 121 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATGATGATTCGTAACAA 180  
Db 464 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATGATGATTCGTAACAA 523  
QY 181 GAATTAGTCGATTTGGTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGTATT 240  
Db 524 GAATTAGTCGATTTGGTTCCCAACACGGTTGTCATGGTGATACCATTCACGTGTATT 583  
QY 241 GAATACATCAACATAATGTGTGTCGTCACAAAGACTACTATCGATACGTTGACGAGAA 300  
Db 584 GAATACATCAACATAATGTGTGTCGTCACAAAGACTACTATCGATACGTTGACGAGAA 643

Qy	301	CAATCATGCCGACCAACAAATGCAACGTTTCGGTATCTCAAACTATTGCAAAATTTAC	360
Db	644	CAATCATGCCGACCAACAAATGCAACGTTTCGGTATCTCAAACTATTGCAAAATTTAC	703
Qy	361	CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC	420
Db	704	CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC	763
Qy	421	ATTATTGGCATCAAAAGATTTAGACGCATTCGGTCATTATGATGCGCGAACCAATCATTCAA	480
Db	764	ATTATTGGCATCAAAAGATTTAGACGCATTCGGTCATTATGATGCGCGAACCAATCATTCAA	823
Qy	481	CGCGATAATGGTTACCAACCAAACTATCACGCTGTCAACATTTGGTTACAGTAACGCA	540
Db	824	CGCGATAATGGTTACCAACCAAACTATCACGCTGTCAACATTTGGTTACAGTAACGCA	883
Qy	541	CAAGGTGTCGATTATTGGATCGTACGAAACAGTTGGGTACCAATTCGGGTGATAATGGT	600
Db	884	CAAGGTGTCGATTATTGGATCGTACGAAACAGTTGGGTACCAATTCGGGTGATAATGGT	943
Qy	601	TACGGTTATTTTGTGCTGCCAACATCGATTTGATGATGATTGAAGAAATATCCATATGTTGTC	660
Db	944	TACGGTTATTTTGTGCTGCCAACATCGATTTGATGATGATTGAAGAAATATCCATATGTTGTC	1003
Qy	661	ATTCTC 666	
Db	1004	ATTCTC 1009	

Search completed: October 28, 2004, 22:43:29  
Job time : 829.956 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 23:44:08 ; Search time 6109.36 Seconds  
(without alignments)  
558.975 Million cell updates/sec

Title: US-09-867-159A-1

Perfect score: 666  
Sequence: 1 actaacgctgcagtatcaaa.....atccatattgtgtcattctc 666

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	666	10	US-09-867-159A-1
2	661.2	99.3	666	15	US-10-001-245-87
3	661.2	99.3	909	15	US-10-297-563-15
4	640.4	96.2	666	15	US-10-001-245-15
5	640.4	96.2	666	15	US-10-001-245-21
6	638.8	95.9	666	15	US-10-001-245-17
7	637.2	95.7	666	15	US-10-001-245-13
8	637.2	95.7	666	15	US-10-001-245-19
9	637.2	95.7	666	15	US-10-001-245-23
10	637.2	95.7	666	15	US-10-001-245-25
11	629.2	94.5	666	15	US-10-001-245-29
12	626	94.0	666	15	US-10-001-245-33

13	624.4	93.8	666	15	US-10-001-245-27	Sequence 27, Appl
14	621.2	93.3	666	15	US-10-001-245-31	Sequence 31, Appl
15	74.4	11.2	100	15	US-10-297-563-8	Sequence 8, Appl
16	63.2	9.5	100	15	US-10-297-563-6	Sequence 6, Appl
17	61.6	9.2	1230	16	US-10-425-114-10408	Sequence 10408, A
18	60.2	9.1	91	15	US-10-297-563-11	Sequence 11, Appl
19	57.2	8.6	95	15	US-10-297-563-10	Sequence 10, Appl
20	56.4	8.5	2167	14	US-10-219-220-102	Sequence 102, App
21	55.8	8.4	1086	9	US-09-938-842A-2484	Sequence 2484, Ap
22	55.8	8.4	1086	11	US-09-938-842A-2484	Sequence 2484, Ap
23	55.6	8.3	322	14	US-10-219-220-113	Sequence 113, App
24	55.4	8.3	741	16	US-10-424-599-118229	Sequence 118229,
25	54	8.1	1186	16	US-10-425-114-13314	Sequence 13314, A
26	54	8.1	1189	16	US-10-424-599-92356	Sequence 92356, A
27	53.2	8.0	1133	15	US-10-369-493-30529	Sequence 30529, A
28	52.2	7.8	82	15	US-10-297-563-13	Sequence 13, Appl
29	51.8	7.7	87	15	US-10-297-563-9	Sequence 9, Appl
30	51.6	7.7	1819	14	US-10-219-220-243	Sequence 243, App
31	50.8	7.6	984	16	US-10-425-114-7415	Sequence 7415, Ap
32	50.6	7.6	1474	14	US-10-219-220-101	Sequence 101, App
33	50	7.5	1356	16	US-10-424-599-127914	Sequence 127914,
34	49.6	7.4	1344	16	US-10-425-114-10837	Sequence 10837, A
35	49.6	7.4	1554	16	US-10-424-599-53151	Sequence 53151, A
36	49.4	7.4	1095	9	US-09-938-842A-2351	Sequence 2351, Ap
37	49.4	7.4	1095	11	US-09-938-842A-2351	Sequence 2351, Ap
38	49	7.4	1392	9	US-09-938-842A-1372	Sequence 1372, Ap
39	49	7.4	1392	11	US-09-938-842A-1372	Sequence 1372, Ap
40	48.4	7.3	1122	9	US-09-938-842A-1324	Sequence 1324, Ap
41	48.4	7.3	1122	11	US-09-938-842A-1324	Sequence 1324, Ap
42	48.4	7.3	1656	16	US-10-425-114-32817	Sequence 32817, A
43	48	7.2	1390	16	US-10-425-114-18625	Sequence 18625, A
44	47.8	7.2	622	17	US-10-767-795-3822	Sequence 3822, Ap
45	47.2	7.1	718	16	US-10-341-961A-155	Sequence 155, App

ALIGNMENTS

RESULT 1  
US-09-867-159A-1  
; Sequence 1, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALLIS  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE REFERENCE: B112812US-antiallis  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 1  
; LENGTH: 666  
; TYPE: DNA  
; ORGANISM: Dermatophagoides pteronyssinus  
US-09-867-159A-1

TERRASSE, GAETAN  
LORIA, EMILE  
TREHIN, YVES

anti-histamine compound

Query Match	100.0%	Score 666	DB 10	Length 666
Best Local Similarity	100.0%	Pred. No. 2.5e-191		
Matches 666	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ACTAACGCTGCAGTATCAATGCAATGCTCCAGCTGCAATCGATTGCGACAAATGCGA	60	
Db	1	ACTAACGCTGCAGTATCAATGCAATGCTCCAGCTGCAATCGATTGCGACAAATGCGA	60	
Qy	61	ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTCTCTGTTGT	120	
Db	61	ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTCTCTGTTGT	120	

[illegible]

```

RESULT 2
US-10-001-245-87
; Sequence 87, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: NO. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 87
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(666)
; OTHER INFORMATION:
US-10-001-245-87

```

Query Match	99.3%;	Score 661.2;	DB 15;	Length 666;
Best Local Similarity	99.5%;	Pred. No. 7.1e-190;		
Matches 663;	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

Qy	1	ACTAAGCGCTGGAGTATCAATGGAAATGCTCAGCTGAAATCGATTGGCAAAATGGCA	60
Db	1	ACTAAGCGCTGCAGTATCAATGGAAATGCTCAGCTGAAATCGATTGGCAAAATGGCA	60
Qy	61	ACTGTCACCTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTTCCTCGTGTT	120
Db	61	ACTGTCACCTCCCATTCGTATGCAAGGAGGCTGTGGTTCATGTTGGGCTTTCCTCGTGTT	120
Qy	121	GCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGATCTTGCCTGAACAA	180
Db	121	GCGCAACTGAATCAGCTTATTTGGCTTACCGTAATCAATCATTTGATCTTGCCTGAACAA	180
Qy	181	GAATTAGTCGATTTGCTTCCCAACACGGTTGTCATGGTGATACCAATCCACGTGGTATT	240
Db	181	GAATTAGTCGATTTGCTTCCCAACACGGTTGTCATGGTGATACCAATCCACGTGGTATT	240
Qy	241	GAATACATCCAAACATAATGTTGCTCCAAAGAAAGCTACTATCGATACGTTGCAACGAGAA	300
Db	241	GAATACATCCAAACATAATGTTGCTCCAAAGAAAGCTACTATCGATACGTTGCAACGAGAA	300
Qy	301	CAATCATCGCGACACACAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC	360
Db	301	CAATCATCGCGACACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC	360
Qy	361	CCACCAAAATGCAAAACAAAATTCGTGAAGCTTTGGCTCAAA	420
Db	361	CCACCAAAATGTAAACAAAATTCGTGAAGCTTTGGCTCAAA	420
Qy	421	ATTATTGCGNTCAAGATTTAGACGCATTCGGTCATTTATGATGCGCGAACAATCATTTCAA	480
Db	421	ATTATTGCGCATCAAGATTTAGACGCATTCGGTCATTTATGATGCGCGAACAATCATTTCAA	480
Qy	481	CCGCGATAATGGTTACCAACAAACTATCACGCTGTCAACATTTGGTTACAGTAACGCA	540
Db	481	CCGCGATAATGGTTACCAACAAACTATCACGCTGTCAACATTTGGTTACAGTAACGCA	540
Qy	541	CAAGGTGTCGATTTATGGATCGTHACGAAACAGTTGGGATACCAATTTGGGGTGNTATGTT	600
Db	541	CAAGGTGTCGATTTATGGATCGTACGAAACAGTTGGGATACCAATTTGGGGTGNTATGTT	600
Qy	601	TACGGTTATTTTGTGCGCAACATCGATTTGATGATGATTCGAAGAAATCCATATGTTGTTC	660
Db	601	TACGGTTATTTTGTGCGCAACATCGATTTGATGATGATTCGAAGAAATCCATATGTTGTTC	660
Qy	661	ATTCTC	666
Db	661	ATTCTC	666

```

RESULT 3
US-10-297-563-15
; Sequence 15, Application US/10297563
; Publication No. US20030138441A1
; GENERAL INFORMATION:
; APPLICANT: Alex Bollen
; APPLICANT: Paul Jacobs
; APPLICANT: Alain Jaquet
; APPLICANT: Marc Georges Francis Massar
; TITLE OF INVENTION: Codon Optimised Recombinant
; TITLE OF INVENTION: Derrnaphagoides Allergens
; FILE REFERENCE: B45219
; CURRENT APPLICATION NUMBER: US/10/297.563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/SP01/06483
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0014288.5
; PRIOR FILING DATE: 2000-06-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 909
; TYPE: DNA

```

ORGANISM: Dermaphagoides pteronyssinus  
US-10-297-563-15

Query Match 99.3%; Score 661.2; DB 15; Length 909;  
Best Local Similarity 99.5%; Pred. No. 8.5e-190;  
Matches 663; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ACTAAGCGCTGAGTATCAATGGAATGCTCAGCTGAAATCGAATTCGACAAATGCGA 60
DB 241 ACTAAGCGCTGAGTATCAATGGAATGCTCAGCTGAAATCGAATTCGACAAATGCGA 300
QY 61 ACTGTCACTCCCATTCGATATGCAAGAGGCTGTGGTTCATGTGGCTTCTCTGGTGT 120
DB 301 ACTGTCACTCCCATTCGATATGCAAGAGGCTGTGGTTCATGTGGCTTCTCTGGTGT 360
QY 121 GCCGAACATGAATCAGCTTATTTGGCTCACGGTAATCAATCAATTCGATTCGTAACAA 180
DB 361 GCCGAACATGAATCAGCTTATTTGGCTCACGGTAATCAATCAATTCGATTCGTAACAA 420
QY 181 GAATTAGTCGATGTGCTTCCCAACACGGTGTGATGCTGATACCAATTCACGTTGATT 240
DB 421 GAATTAGTCGATGTGCTTCCCAACACGGTGTGATGCTGATACCAATTCACGTTGATT 480
QY 241 GAATACATCCAAATATAGTGTCTGCCAAGAAAGCTACTATCGATAGCTTGCACGAGAA 300
DB 481 GAATACATCCAAATATAGTGTCTGCCAAGAAAGCTACTATCGATAGCTTGCACGAGAA 540
QY 301 GAATATGCGGACACCAAAATGCAAGCTTTCGGTATCTCAAACTATTCGCAATTTAC 360
DB 541 GAATATGCGGACACCAAAATGCAAGCTTTCGGTATCTCAAACTATTCGCAATTTAC 600
QY 361 CCACCAAAATGCAAAACAAATTCGTGAAGCTTGGCTCAAAACACAGCGTATTCGGCTC 420
DB 601 CCACCAAAATGCAAAACAAATTCGTGAAGCTTGGCTCAAAACACAGCGTATTCGGCTC 660
QY 421 ATTATTTGGCATCAAAAGATTTAGACGATTCCTGATATTAATGATGCGGAAACAAATCAA 480
DB 661 ATTATTTGGCATCAAAAGATTTAGACGATTCCTGATATTAATGATGCGGAAACAAATCAA 720
QY 481 CCGGATATGTTTACCAACCAATATCAGCTGTCACATTTGTTGTTAGTACAGTAACGCA 540
DB 721 CCGGATATGTTTACCAACCAATATCAGCTGTCACATTTGTTGTTAGTACAGTAACGCA 780
QY 541 CAAGGTGTCGATTAATTTGATGCTAGCAACAGTCAGTGGGATACCAATTTGGGTTGATAATG 600
DB 781 CAAGGTGTCGATTAATTTGATGCTAGCAACAGTCAGTGGGATACCAATTTGGGTTGATAATG 840
QY 601 TACGGTATTTGCTGCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 841 TACGGTATTTGCTGCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 661 ATTCTC 666
DB 901 ATTCTC 906
```

RESULT 4  
US-10-001-245-15  
; Sequence 15, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1e1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16

NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn version 3.1

SEQ ID NO 15  
LENGTH: 666  
TYPE: DNA  
ORGANISM: Dermaphagoides pteronyssinus  
FEATURE:

```
NAME/KEY: CDS
LOCATION: (1)..(666)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (70)..(72)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (232)..(234)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (148)..(150)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (295)..(297)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (466)..(468)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (481)..(483)
OTHER INFORMATION:
FEATURE:
NAME/KEY: mutation
LOCATION: (499)..(501)
OTHER INFORMATION:
```

US-10-001-245-15

Query Match 96.2%; Score 640.4; DB 15; Length 666;

Best Local Similarity 97.6%; Pred. No. 1.5e-183;  
Matches 650; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
QY 1 ACTAAGCGCTGAGTATCAATGGAATGCTCAGCTGAAATCGAATTCGACAAATGCGA 60
DB 1 ACTAAGCGCTGAGTATCAATGGAATGCTCAGCTGAAATCGAATTCGACAAATGCGA 60
QY 61 ACTGTCACTCCCATTCGATATGCAAGAGGCTGTGGTTCATGTGGCTTCTCTGGTGT 120
DB 61 ACTGTCACTCCCATTCGATATGCAAGAGGCTGTGGTTCATGTGGCTTCTCTGGTGT 120
QY 121 GCCGAACATGAATCAGCTTATTTGGCTCACGGTAATCAATCAATTCGATTCGTAACAA 180
DB 121 GCCGAACATGAATCAGCTTATTTGGCTCACGGTAATCAATCAATTCGATTCGTAACAA 180
QY 181 GAATTAGTCGATGTGCTTCCCAACACGGTGTGATGCTGATACCAATTCACAGGTTAT 240
DB 181 GAATTAGTCGATGTGCTTCCCAACACGGTGTGATGCTGATACCAATTCACAGGTTAT 240
QY 241 GAATACATCCAAATATAGTGTCTGCCAAGAAAGCTACTATCGATAGCTTGCACGAGAA 300
DB 241 GAATACATCCAAATATAGTGTCTGCCAAGAAAGCTACTATCGATAGCTTGCACGAGAA 300
QY 301 CAATCATGCGGACGACCAAAATTCGTAAGCTTTCGGTATCTCAAACTATTCGCAAAATTTAC 360
DB 301 CAATCATGCGGACGACCAAAATTCGTAAGCTTTCGGTATCTCAAACTATTCGCAAAATTTAC 360
QY 361 CCACCAAAATGCAAAACAAATTCGTAAGCTTTCGGTATCTCAAACTATTCGCAAAATTTAC 420
DB 361 CCACCAAAATGCAAAACAAATTCGTAAGCTTTCGGTATCTCAAACTATTCGCAAAATTTAC 420
QY 421 ATTATTTGGCATCAAAAGATTTAGACGATTCCTGATATTAATGATGCGGAAACAAATCAA 480
```

421 ATTATTGGCATCAAGATTTAGACGCAATTCGGTCATTATGATGGCGAAACAATCATTTCAA 480  
481 CGCGATAATGGTTTACCAACCAAACTATCAGCTGTCAACATTTGTTGGTTACAGTAACGCA 540  
481 CAGATAAATGGTTTACCAACCAAACTATCAGCTGTCAACATTTGTTGGTTACAGTAACGCA 540  
541 CAAGGTGTGATATTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTGATAATGTT 600  
541 CAAGGTGTGATATTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTGATAATGTT 600  
601 TAGGGTATTCTCTGCCACATCGATTTGATGATGATTGAGATATCCATATGTTGTC 660  
601 TAGGGTATTCTCTGCCACATCGATTTGATGATGATTGAGATATCCATATGTTGTC 660  
661 ATTCTC 666  
661 ATTCTC 666

RESULT 5  
US-10-001-245-21  
; Sequence 21, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 21  
; LENGTH: 666  
; TYPE: DNA  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(666)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (70)..(72)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (148)..(150)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (199)..(201)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (295)..(297)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (466)..(468)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (481)..(483)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation

; LOCATION: (499)..(501)  
; OTHER INFORMATION:  
US-10-001-245-21  
Query Match 96.2%; Score 640.4; DB 15; Length 666;  
Best Local Similarity 97.6%; Pred. No. 1.5e-183;  
Matches 650; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
QY 1 ACTAACCCCTGCAGTATCAATGGAAATGCTCCAGCTGAAATCGAATTCGACAAATCGGA 60  
Db 1 ACTAACCCCTGCAGTATCAATGGAAATGCTCCAGCTGAAATCGAATTCGACAAATCGGA 60  
QY 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGCTTCAATGTTGGGCTTCTCTGTGTT 120  
Db 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGTGCTTCAATGTTGGGCTTCTCTGTGTT 120  
QY 121 GCCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGTGAACA 180  
Db 121 GCCGCAACTGAATCAGCTTATTTGGCTCACCGTAATCAATCATTTGGATCTTGTGAACA 180  
QY 181 GAATTAGTCGATTCGCTTCCCAACACGTTGTGTCGATGATACCATTCACCTGGTATT 240  
Db 181 GAATTAGTCGATTCGCTTCCCAACACGTTGTGTCGATGATACCATTCACCTGGTATT 240  
QY 241 GAATACATCCAAACATAATGGTGTCTCCAAAGAAAGCTTACTATCGATACGTTGCACGAA 300  
Db 241 GAATACATCCAAACATAATGGTGTCTCCAAAGAAAGCTTACTATCGATACGTTGCACGAA 300  
QY 301 CAATCATGCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 301 CAATCATGCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
QY 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
Db 361 CCACCAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420  
QY 421 ATTATGCGATCAAAAGATTAGACGANTTCGGTCAATTCGATGCGCGAACAATCATTTCAA 480  
Db 421 ATTATGCGATCAAAAGATTAGACGANTTCGGTCAATTCGATGCGCGAACAATCATTTCAA 480  
QY 481 CGCGATATGTTTACCAACCAAACTATCAGCTGTCAACATTTGTTACAGTAACGCA 540  
Db 481 CGCGATATGTTTACCAACCAAACTATCAGCTGTCAACATTTGTTACAGTAACGCA 540  
QY 541 CAAGGTGTGATATTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTGATAATGTT 600  
Db 541 CAAGGTGTGATATTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTGATAATGTT 600  
QY 601 TAGGGTATTCTCTGCCACATCGATTTGATGATGATTGAGATATCCATATGTTGTC 660  
Db 601 TAGGGTATTCTCTGCCACATCGATTTGATGATGATTGAGATATCCATATGTTGTC 660  
QY 661 ATTCTC 666  
Db 661 ATTCTC 666

RESULT 6  
US-10-001-245-17  
; Sequence 17, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361



Qy	421	ATTATTGGCATCAAGATTTTAGACGATTCCCGTCAATTATGATGCCGGAACAATCATTCATCA	480
Db	421	ATTATTGGCATCAAGATTTTAGACGATTCCCGTCAATTATGATGCCGGAACAATCATTCATCA	480
Qy	481	CGCGATAATGGTTACCAACCAAACTATCACGGTGTCAACATTGTTGGTTACAGTAACGCA	540
Db	481	CAGGATTAATGGTTACCAACCAAACTATCACGGTGTCAACATTGTTGGTTACAGTAACGCA	540
Qy	541	CAAGGTGTCGATTATTGGATCGTACGAACACAGTTGGGATACCAATTTGGGGTGATAATGGT	600
Db	541	CAAGGTGTCGATTATTGGATCGTACGAACACAGTTGGGATACCAATTTGGGGTGATAATGGT	600
Qy	601	TACGGTTATTTTGCTGCCAACATCGGATTTTGATGATTGAAGAATATCCATATGTTGTC	660
Db	601	TACGGTTATTTTGCTGCCAACATCGGATTTTGATGATTGAAGAATATCCATATGTTGTC	660
Qy	661	ATTCTC	666
Db	661	ATTCTC	666

```

RESULT 7
US-10-001-245-13
; Sequence 13, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFØRT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(666)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (70)..(72)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (148)..(150)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (232)..(234)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (295)..(297)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (466)..(468)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (481)..(483)
; OTHER INFORMATION:
; FEATURE:

```

```

; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(666)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (148)..(150)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (232)..(234)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (295)..(297)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (466)..(468)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (481)..(483)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (499)..(501)
; OTHER INFORMATION:
; US-10-001-245-17

Query Match          95.9%; Score 638.8; DB 15; Length 666;
Best Local Similarity 97.4%; Pred. No. 4.5e-103;
Matches 649; Conservative 0; Mismatches 17; Indels 0; Gaps 0

QY      1 ACTACGGCTGCAGTATCAATGGAAATGCTCCAGCTGAAATCGAATTGGCAAAATGCGA 60
        |||||
Db       1 ACTACGGCTGCAGTATCAATGGAAATGCTCCAGCTGAAATCGAATTGGCAAAATGCGA 60

QY      61 ACTGTCACTCCCATTCGATATCGAAGGAGGCTGTGGTTCATGTTGGCTTTCTCTGGTGT 120
        |||||
Db       61 ACTGTCACTCCCATTCGATATCGAAGGAGGCTGTGGTTCATGTTGGCTTTCTCTGGTGT 120

QY      121 GCGCAACTGAATCAGCTATTATTGGCTCACGGTAATCAATCATATGATCTTGCTGAACAA 180
        |||||
Db       121 GCGCAACTGAATCAGCTATTATTGGCTCACGGTAATCAATCATATGATCTTGCTGAACAA 180

QY      181 GAATTAGTCGATTGTGCTTTCCCAACACGGTTTGCATGTCATACCAATTCACGTTGTTAT 240
        |||||
Db       181 GAATTAGTCGATTGTGCTTTCCCAACACGGTTTGCATGTCATACCAATTCACGTTGTTAT 240

QY      241 GAATACATCCAAACATAATGGTGTGTCGTCGAAGAAAGCTACTATCGATACGTTCCACAGAA 300
        |||||
Db       241 GAATACATCCAAACATAATGGTGTGTCGTCGAAGAAAGCTACTATCGATACGTTCCACAGAA 300

QY      301 CAATCATGCCGACCAACCACAAATGCCAAGTTTTCGGTATCTCAAACATTATGCCAAATTTAC 360
        |||||
Db       301 CAATCATGCCGACCAACCACAAATGCCAAGTTTTCGGTATCTCAAACATTATGCCAAATTTAC 360

QY      361 CCACCAATTCGAAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
        |||||
Db       361 CCACCAATTCGAAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420

```

NAME/KEY: mutation  
LOCATION: (508)..(510)  
OTHER INFORMATION:  
US-10-001-245-13

Query Match 95.7%; Score 637.2; DB 15; Length 666;  
Best Local Similarity 97.3%; Pred. No. 1.4e-182;  
Matches 648; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY	1	ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA	60
DB	1	ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA	60
QY	61	ACTGTCACTCCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGCTTCTCTGGTGT	120
DB	61	ACTGTCACTCCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGCTTCTCTGGTGT	120
QY	121	GCGCAACTGAATCAGCTTATTTGGCTCAACGCTGAAATCAATTCATTCGATTCGTCGAACAA	180
DB	121	GCGCAACTGAATCAGCTTATTTGGCTCAACGCTGAAATCAATTCATTCGATTCGTCGAACAA	180
QY	181	GAATTAGTCGATTCGCTCCCAACACGCTGTCATGCTGATACCATTCACGCTGATTT	240
DB	181	GAATTAGTCGATTCGCTCCCAACACGCTGTCATGCTGATACCATTCACGCTGATTT	240
QY	241	GAATACATCCAAACATATGCTGCTCCCAAGAAAGCTACTATCGATACGTTGCGACAGAA	300
DB	241	GAATACATCCAAACATATGCTGCTCCCAAGAAAGCTACTATCGATACGTTGCGACAGAA	300
QY	301	CAATCATGCGGACACCAAAATGCAACAGTTGCGTATCTCAAACTATTCGCAAAATTTAC	360
DB	301	CAATCATGCGGACACCAAAATGCAACAGTTGCGTATCTCAAACTATTCGCAAAATTTAC	360
QY	361	CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTCGCGTC	420
DB	361	CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTCGCGTC	420
QY	421	ATTATTGGCATCAAGATTTAGACGATTCGCTCATATGATGGCGCAACATTCATCAA	480
DB	421	ATTATTGGCATCAAGATTTAGACGATTCGCTCATATGATGGCGCAACATTCATCAA	480
QY	481	CGGATATGTTTACCAACAACTATCAGCTGTCACATTCGTTGCTTACAGTACGCA	540
DB	481	GAGATATGTTTACCAACAACTATCAGCTGTCACATTCGTTGCTTACAGTACGCA	540
QY	541	CAAGGTGCGATTTATTCGATTCGTAACAAAGTTGGATACCAATTCGGGTGATATGT	600
DB	541	CAAGGTGCGATTTATTCGATTCGTAACAAAGTTGGATACCAATTCGGGTGATATGT	600
QY	601	TACGTTATTTGCTGCAACATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	660
DB	601	TACGTTATTTGCTGCAACATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG	660
QY	661	ATTCTC 666	
DB	661	ATTCTC 666	

## RESULT 8

US-10-001-245-19  
Sequence 19, Application US/10001245  
Publication No. US20030175312A1  
GENERAL INFORMATION:  
APPLICANT: HOLM, Jens  
APPLICANT: IPSEN, Henrik  
APPLICANT: LARSEN, Jorgen N.  
APPLICANT: SPANGFORT, Michael D.  
TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
FILE REFERENCE: 4305/1H942-US2  
CURRENT APPLICATION NUMBER: US/10/001.245  
CURRENT FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/298,170  
PRIOR FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/249,361  
PRIOR FILING DATE: 2000-11-16  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 19  
LENGTH: 666  
TYPE: DNA  
ORGANISM: Dermatophagoides pteronyssinus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(666)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (58)..(60)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (232)..(234)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (148)..(150)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (223)..(234)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (295)..(297)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (468)..(468)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (481)..(483)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (499)..(501)  
OTHER INFORMATION:  
US-10-001-245-19

Query Match 95.7%; Score 637.2; DB 15; Length 666;  
Best Local Similarity 97.3%; Pred. No. 1.4e-182;  
Matches 648; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY	1	ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA	60
DB	1	ACTAACGCTGCGATGATCAATGGAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA	60
QY	61	ACTGTCACTCCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGCTTCTCTGGTGT	120
DB	61	ACTGTCACTCCCATTCGTATGCAAGAGGCTGGTTTCATGTTGGCTTCTCTGGTGT	120
QY	121	GCGCAACTGAATCAGCTTATTTGGCTCAACGCTGAAATCAATTCATTCGATTCGTCGAACAA	180
DB	121	GCGCAACTGAATCAGCTTATTTGGCTCAACGCTGAAATCAATTCATTCGATTCGTCGAACAA	180
QY	181	GAATTAGTCGATTCGCTCCCAACACGCTGTCATGCTGATACCATTCACGCTGATTT	240
DB	181	GAATTAGTCGATTCGCTCCCAACACGCTGTCATGCTGATACCATTCACGCTGATTT	240
QY	241	GAATACATCCAAACATATGCTGCTCCCAAGAAAGCTACTATCGATACGTTGCGACAGAA	300
DB	241	GAATACATCCAAACATATGCTGCTCCCAAGAAAGCTACTATCGATACGTTGCGACAGAA	300
QY	301	CAATCATGCGGACACCAAAATGCAACAGTTGCGTATCTCAAACTATTCGCAAAATTTAC	360
DB	301	CAATCATGCGGACACCAAAATGCAACAGTTGCGTATCTCAAACTATTCGCAAAATTTAC	360
QY	361	CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTCGCGTC	420
DB	361	CCACCAATGCAACAAATTCGTGAAGCTTTGGCTCAACCCACACGCGTATTCGCGTC	420

QY 421 ATTATTGGCATCAAGATTAGACGCAATTCGGTCAATATGATGGCCGCAACATCAATTCAA 480  
Db 421 ATTATTGGCATCAAGATTAGACGCAATTCGGTCAATATGATGGCCGCAACATCAATTCAA 480  
QY 481 CCGGATAATGGTTACCAACCAACTATACCGTGTCAACATTTGGTTACAGTAACGCA 540  
Db 481 CAGGATAATGGTTACCAACCAACTATACCGTGTCAACATTTGGTTACAGTAACGCA 540  
QY 541 CAGGTGTGATATTGGATCTAGCAACAGTTGGGATACCAATTTGGGTTGATAATGGT 600  
Db 541 CAGGTGTGATATTGGATCTAGCAACAGTTGGGATACCAATTTGGGTTGATAATGGT 600  
QY 601 TACGGTTATTTTGGTCCCAACATCGATTTGATGATGATTTGAAGATATCCATATGTTGC 660  
Db 601 TACGGTTATTTTGGTCCCAACATCGATTTGATGATGATTTGAAGATATCCATATGTTGC 660  
QY 661 ATTCTC 666  
Db 661 ATTCTC 666

RESULT 9  
US-10-001-245-23  
; Sequence 23, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 23  
; LENGTH: 666  
; TYPE: DNA  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(666)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (58)..(60)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (148)..(150)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (199)..(201)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (295)..(297)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (466)..(468)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (481)..(483)  
; OTHER INFORMATION:

FEATURE:  
; NAME/KEY: mutation  
; LOCATION: (499)..(501)  
; OTHER INFORMATION:  
US-10-001-245-23

Query Match  
Best Local Similarity 95.7%; Score 637.2; DB 15; Length 666;  
Best Local Similarity 97.3%; Pred. No. 1.4e-182;  
Matches 648; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTTGGCAGCAAAATCGCA 60  
Db 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAATCGATTTGGCAGCAAAATCGCA 60  
QY 61 ACTGTCACTCCCATTCGTATGCAAGAGGCTGTGGTTTCATGTTGGGCTTTCTCTGGTGT 120  
Db 61 ACTGTCACTCCCATTCGTATGCAAGAGGCTGTGGTTTCATGTTGGGCTTTCTCTGGTGT 120  
QY 121 GCAGCAACTGAATCAGCTTATTTGGTCTCACCGTAACTCAATCATTCATGATCTGCTGAACAA 180  
Db 121 GCAGCAACTGAATCAGCTTATTTGGTCTCACCGTAACTCAATCATTCATGATCTGCTGAACAA 180  
QY 181 GAATTAGTCTGATTCGCTCCCAACACGGTTGTCATGCTGATACCATTCACCGTGTATT 240  
Db 181 GAATTAGTCTGATTCGCTCCCAACACGGTTGTCATGCTGATACCATTCACCGTGTATT 240  
QY 241 GAATACATCCAAACATATGCTGCTCCAAAGAGCTACTATCGATACGTTGCAACGAA 300  
Db 241 GAATACATCCAAACATATGCTGCTCCAAAGAGCTACTATCGATACGTTGCAACGAA 300  
QY 301 CATCATCGCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 301 CATCATCGCGACCAACCAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
QY 361 CCACCAAAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCGGTC 420  
Db 361 CCACCAAAATGCAACCAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCGGTC 420  
QY 421 ATTATTCGCATCAAGATTTAGACGCAATTCGGTCAATTCATGATGCGCGCAACATCAATCAA 480  
Db 421 ATTATTCGCATCAAGATTTAGACGCAATTCGGTCAATTCATGATGCGCGCAACATCAATCAA 480  
QY 481 CGCGATAATGGTTACCAACCAAACTACCGCTGTCAAACTATTGGTTACAGTAACGCA 540  
Db 481 GAAGTAAATGGTTACCAACCAAACTACCGCTGTCAAACTATTGGTTACAGTAACGCA 540  
QY 541 CAAGGTGTGATTTGGTGTGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATAATGGT 600  
Db 541 CAAGGTGTGATTTGGTGTGATCGTACGAAACAGTTGGGATACCAATTTGGGTTGATAATGGT 600  
QY 601 TACGGTTATTTGCTGCGCAACATCGATTTGATGATGATTTGAAGATATCCATATGTTGC 660  
Db 601 TACGGTTATTTGCTGCGCAACATCGATTTGATGATGATTTGAAGATATCCATATGTTGC 660  
QY 661 ATTCTC 666  
Db 661 ATTCTC 666

RESULT 10  
US-10-001-245-25  
; Sequence 25, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170

```

; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 25
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(666)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (58)..(60)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (148)..(150)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (199)..(201)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (295)..(297)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (456)..(458)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (481)..(483)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (499)..(501)
; OTHER INFORMATION:
; US-10-001-245-25

Query Match          95.7%; Score 637.2; DB 15; Length 666;
Best Local Similarity 97.3%; Pred. No. 1.4e-182;
Matches 648; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCACAATGCGA 60
Db 1 ACTAACCCCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCACAATGCGA 60
Qy 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGGTTCATGTTGGGCTTTCTCTGGTGT 120
Db 61 ACTGTCACTCCCATTCGTATGCAAGGAGGCTGGTTCATGTTGGGCTTTCTCTGGTGT 120
Qy 121 GCGCAACTGAATCAGCTATTTGGCTCACCGGTAATCAATCATTTGGATCTTCTGCAACA 180
Db 121 GCGCAACTGAATCAGCTATTTGGCTGCGGTAATCAATCATTTGGATCTTCTGCAACA 180
Qy 181 GAATTAGTCGATTTGCTTCCCAACACGGTGTGTCATGGTGCATACCAATCCAGTGGTAT 240
Db 181 GAATTAGTCGATTTGCTTCCCAACACGGTGTGTCATGGTGCATACCAATCCAGTGGTAT 240
Qy 241 GAATACATCCACATATGCTGTCGTCGAAGAAGCTACTATCGATACGTTGACCGAGAA 300
Db 241 GAATACATCCACATATGCTGTCGTCGAAGAAGCTACTATCGATACGTTGACCGAGAA 300
Qy 301 CAATCATGCGCAGCACAATGCAACAGTTTCGGTATCTCAAACTATTCGCAAAATTTAC 360
Db 301 CAATCATGCGCAGCACAATGCAACAGTTTCGGTATCTCAAACTATTCGCAAAATTTAC 360
Qy 361 CCACCAAAATGCAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACACGCGCTATTCGCGTC 420
Db 361 CCACCAAAATGCAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACACGCGCTATTCGCGTC 420
Qy 421 ATTATTGGCATCAAAAGATTTAGACGCAATTCGGTCATTATGATGCGCGAACAATCAATCAA 480
Db 421 ATTATTGGCATCAAAAGATTTAGACGCAATTCGGTCATTATGATGCGCGAACAATCAATCAA 480
Qy 481 CGGATAATGTTTACCAACCAAACTATCACGCTGTCAACATTTGTTTACAGTAACGCA 540
Db 481 GAAGATAATGTTTACCAACCAAACTATCACGCTGTCAACATTTGTTTACAGTAACGCA 540
Qy 541 CAAGTGTCGATTTGATGCTAGCAAGACAGTGGGATACCAATTCGGGTGATATGGT 600
Db 541 CAAGTGTCGATTTGATGCTAGCAAGACAGTGGGATACCAATTCGGGTGATATGGT 600
Qy 601 TACGGTTATTTTGTGTCACACATCGATTTGATGATGATTTGAAGAATATCCATATGTTGTC 660
Db 601 TACGGTTATTTTGTGTCACACATCGATTTGATGATGATTTGAAGAATATCCATATGTTGTC 660
Qy 661 ATTCTC 666
Db 661 ATTCTC 666.

RESULT 11
US-10-001-245-29
; Sequence 29, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 29
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(666)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (37)..(39)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (70)..(72)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (148)..(150)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (232)..(234)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (295)..(297)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (325)..(327)
```

OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (382)..(384)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (466)..(468)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (481)..(483)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (499)..(501)  
OTHER INFORMATION:  
US-10-001-245-29

Query Match 94.5%; Score 629.2; DB 15; Length 666;  
Best Local Similarity 96.5%; Pred. No. 3.7e-180;  
Matches 643; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY	1	ACTAAGCGCTGAGTATCAATGAAATGCTCCAGCTGAAATCGATTTCGGACAAATGCGA	60
DB	1	ACTAAGCGCTGAGTATCAATGAAATGCTCCAGTAGCATCGATTTCGGACAAATGCGA	60
QY	61	ACTGCACTCCCAATCGTATGCAAGAGGCTGTGGTTCATGTTGGCTTTCTCTGGTGT	120
DB	61	ACTGCACTCAATTCGTATGCAAGAGGCTGTGGTTCATGTTGGCTTTCTCTGGTGT	120
QY	121	GCGCAACTGATCATGCTTATTTGGCTCACCTGATCAATCAATGCTGCTGAAACAA	180
DB	121	GCGCAACTGATCACTTATTTGGCTGCGTATCAATCAATGCTGCTGAAACAA	180
QY	181	GAATTAGTCGATTGCTTCCCAACACGCTTGTCTATGCTGATACCAATCCACGTTGAT	240
DB	181	GAATTAGTCGATTGCTTCCCAACACGCTTGTCTATGCTGATACCAATCCACGTTGAT	240
QY	241	GAATACATCCACATATGCTGCTCCAGAAAGCTACTATCGATAGTTGCACAGAA	300
DB	241	GAATACATCCACATATGCTGCTCCAGAAAGCTACTATCGATAGTTGCACAGAA	300
QY	301	CAATCATGCGGACCAACCAATGCAACCGTTTCGGTATCTCAAACTATTGCCAAATTTAC	360
DB	301	CAATCATGCGGACCAACCAATGCAACCGTTTCGGTATCTCAAACTATTGCCAAATTTAC	360
QY	361	CCACCAATGCAAAACAAATTCGTGAAGCTTTGGCTCAAAACCAAGCGCTATTGCCGTC	420
DB	361	CCACCAATGTAACAAAAATTCAGGAAGCTTTGGCTCAAAACCAAGCGCTATTGCCGTC	420
QY	421	ATTATTGGCATCAAGATTTAGACGCTTCCGTCATTATGATGCGCGCAACATCAATCAA	480
DB	421	ATTATTGGCATCAAGATTTAGACGCTTCCGTCATTATGATGCGCGCAACATCAATCAA	480
QY	481	CAGGATAATGGTTACCAACCACTATACGCTGTCACATTTGGTTACAGTAACGCA	540
DB	481	CAGGATAATGGTTACCAACCACTATACGCTGTCACATTTGGTTACAGTAACGCA	540
QY	541	CAAGGTGTCGATTTGGATTCGTACCAACAGCTGGGATACCAATGGGTCGATATCGT	600
DB	541	CAAGGTGTCGATTTGGATTCGTACCAACAGCTGGGATACCAATTTGGGTCGATATCGT	600
QY	601	TACGGTTATTTCGTCGCCAACATCGATTTCGATGATGATTTGAAGAATATCCATATGTC	660
DB	601	TACGGTTATTTCGTCGCCAACATCGATTTCGATGATGATTTGAAGAATATCCATATGTC	660
QY	661	ATTCTC 666	
DB	661	ATTCTC 666	

RESULT 12

US-10-001-245-33  
Sequence 33, Application US/10001245  
Publication No. US20030175312A1  
GENERAL INFORMATION:  
APPLICANT: HOLM, Jens  
APPLICANT: IPSEN, Henrik  
APPLICANT: LARSEN, Jorgen N.  
APPLICANT: SPANGFORT, Michael D.  
TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
FILE REFERENCE: 4305/1H942-US2  
CURRENT APPLICATION NUMBER: US/10/001,245  
CURRENT FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/298,170  
PRIOR FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/249,361  
PRIOR FILING DATE: 2000-11-16  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 33  
LENGTH: 666  
TYPE: DNA  
ORGANISM: Dermatophagoides pteronyssinus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(666)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (37)..(39)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (70)..(72)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (148)..(150)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (232)..(234)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (295)..(297)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (325)..(327)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (382)..(384)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (466)..(468)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (481)..(483)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (499)..(501)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (574)..(576)  
OTHER INFORMATION:  
US-10-001-245-33

Query Match 94.0%; Score 626; DB 15; Length 666;

Best Local Similarity 96.2%; Pred. No. 3.4e-179; Mismatches 25; Indels 0; Gaps 0;  
Matches 641; Conservative 0;

Qy 1 ACTAACGCTGCAGTATCAATGGAAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60  
Db 1 ACTAACGCTGCAGTATCAATGGAAATGCTCCAGCTAGCATCGAATTCGACAAATGCGA 60

Qy 61 ACTGTCACTCCCATTCGATGCAAGGAGCTGTGGTTCATGTTGGCTTCTCTGGTGT 120  
Db 61 ACTGTCACTCCCATTCGATGCAAGGAGCTGTGGTTCATGTTGGCTTCTCTGGTGT 120

Qy 121 GCCCAACTGAATCAGCTTATTTGGCTCACCGCTCAATCAATCAATCAATCAATCAATCAAT 180  
Db 121 GCCCAACTGAATCAGCTTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

Qy 181 GAATTAGTCGATTCGCTTCCCAACAGGTTGTCATGGTGATACCAATTCACGTGTATT 240  
Db 181 GAATTAGTCGATTCGCTTCCCAACAGGTTGTCATGGTGATACCAATTCACGTGTATT 240

Qy 241 GAATACATCCAAACATAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCAACGAGAA 300  
Db 241 GAATACATCCAAACATAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCAACGAGAA 300

Qy 301 CAATCATGCGGACCAACAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360  
Db 301 CAATCATGCGGACCAACAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360

Qy 361 CCACCAATGCAACAAATTCGTGAGCTTTCGCTCAACCCACAGCGCTATTGCGGTC 420  
Db 361 CCACCAATGCAACAAATTCGTGAGCTTTCGCTCAACCCACAGCGCTATTGCGGTC 420

Qy 421 ATTATGGCATCAAGATTTAGAGCATTCGCTCATTTATGATGCGCGCAACATCAATCAA 480  
Db 421 ATTATGGCATCAAGATTTAGAGCATTCGCTCATTTATGATGCGCGCAACATCAATCAA 480

Qy 481 CGCGATATGTTACCAACCAACTATCAGCTGCTCAACATTCGTTGCTTACATGACGCA 540  
Db 481 CAGGATATGTTTACCAACCAACTATCAGCTGCTCAACATTCGTTGCTTACATGACGCA 540

Qy 541 CAAGGTGTCGATTTATGATGCTGCAAGACAGTTGGGATACCAATTCGGGTGATAATGTT 600  
Db 541 CAAGGTGTCGATTTATGATGCTGCAAGACAGTTGGGATACCAATTCGGGTGATAATGTT 600

Qy 601 TACGGTATTTTGTGCAACATGATTTGATGATGATGATGATGATGATGATGATGATGATG 660  
Db 601 TACGGTATTTTGTGCAACATGATTTGATGATGATGATGATGATGATGATGATGATGATG 660

Qy 661 ATTCTC 666  
Db 661 ATTCTC 666

## RESULT 13

US-10-001-245-27  
; Sequence 27, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 666

TYPE: DNA  
ORGANISM: Dermatophagoides pteronyssinus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(666)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (37)..(39)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (70)..(72)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (148)..(150)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (232)..(234)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (295)..(297)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (325)..(327)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (382)..(384)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (466)..(468)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (481)..(483)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mutation  
LOCATION: (499)..(501)  
OTHER INFORMATION:  
US-10-001-245-27

Query Match 93.8%; Score 624.4; DB 15; Length 666;  
Best Local Similarity 96.1%; Pred. No. 1.1e-178;  
Matches 640; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ACTAACGCTGCAGTATCAATGGAAATGCTCCAGCTGAAATCGAATTCGACAAATGCGA 60  
Db 1 ACTAACGCTGCAGTATCAATGGAAATGCTCCAGCTAGCATCGAATTCGACAAATGCGA 60

Qy 61 ACTGTCACTCCCATTCGATGCAAGGAGCTGTGGTTCATGTTGGCTTCTCTGGTGT 120  
Db 61 ACTGTCACTCCCATTCGATGCAAGGAGCTGTGGTTCATGTTGGCTTCTCTGGTGT 120

Qy 121 GCCCAACTGAATCAGCTTATTTGGCTCACCGTAAATCAATCAATCAATCAATCAATCAAT 180  
Db 121 GCCCAACTGAATCAGCTTATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

Qy 181 GAATTAGTCGATTCGCTTCCCAACAGGTTGTCATGGTGATACCAATTCACGTGTATT 240  
Db 181 GAATTAGTCGATTCGCTTCCCAACAGGTTGTCATGGTGATACCAATTCACGTGTATT 240

Qy 241 GAATACATCCAAACATAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCAACGAGAA 300  
Db 241 GAATACATCCAAACATAATGCTGCTCCAAAGAGCTACTATCGATACGTTGCAACGAGAA 300

Qy 301 CAATCATGCGGACCAACAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360

```
Db 301 CAATCATGCGGACGACCAAAATCGATCGTTTCGGTATCTCAAACTATTGCCAAATTAC 360
Qy 361 CCACCAAAATGCAAAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Db 361 CCACCAAAATGTAACAAAATTCGAAGAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Qy 421 ATTATTGGCATCAAGATTTAGACGCATTCCTGTCATTTATGATGGCGGACAAATCAATCAA 480
Db 421 ATTATTGGCATCAAGATTTAGACGCATTCCTGTCATTTATGATGGCGGACAAATCAATCAA 480
Qy 481 CGCGAATGTTTACCAACCAACTATCAGCTGTCACAACTTTGTTTACAGTAACGCA 540
Db 481 GAAGATAATGTTTACCAACCAACTATCAGCTGTCACAACTTTGTTTACAGTAACGCA 540
Qy 541 CAAAGTGTGCGATTTATTTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTGATTAATGGT 600
Db 541 CAAAGTGTGCGATTTATTTGGATTCGTACGAAACAGTTGGGATACCAATTTGGGGTGATTAATGGT 600
Qy 601 TAGCGTTATTGCTGCCAACATCGATTTTGGATGATGATTAAGAAATATCCATATGTTGTC 660
Db 601 TAGCGTTATTGCTGCCAACATCGATTTTGGATGATGATTAAGAAATATCCATATGTTGTC 660
Qy 661 ATTCTC 666
Db 661 ATTCTC 666
```

## RESULT 14

```
US-10-001-245-31
; Sequence 31, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1e1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(666)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (37)..(39)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (70)..(72)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (148)..(150)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (232)..(234)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
```

```
; LOCATION: (295)..(297)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (325)..(327)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (382)..(384)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (466)..(468)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (481)..(483)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (499)..(501)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (574)..(576)
; OTHER INFORMATION:
US-10-001-245-31
```

Query Match 93.3%; Score 621.2; DB 15; Length 666;

```
Best Local Similarity 95.8%; Pred. No. 9,9e-178;
Matches 638; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
Qy 1 ACTAACGCGCTGCAGTATCAATGGAATGCTCCAGCTGAAATCGATTTGGCGACAAATCGCA 60
Db 1 ACTAACGCGCTGCAGTATCAATGGAATGCTCCAGCTAGCATCGATTTGGCACAATCGCA 60
Qy 61 ACTGTCATCCCATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGTTGTT 120
Db 61 ACTGTCATCCCATTCGTATGCAAGAGGCTGTGTTTCATGTTGGGCTTTCTCTGTTGTT 120
Qy 121 GCCGCAACTGNAATCAGCTTATTGGCTCAAGTAATCAATCAATCAATCAATCAATCAAT 180
Db 121 GCCGCAACTGNAATCAGCTTATTGGCTGTGCGTAATCAATCAATCAATCAATCAATCAAT 180
Qy 181 GAATTAGTCGATTGCTGCTCCCAACAGGTTGTCATGGTGATACCATCCACGCTGTTATT 240
Db 181 GAATTAGTCGATTGCTGCTCCCAACAGGTTGTCATGGTGATACCATCCACGAGGTATT 240
Qy 241 GAATACATCCAACTAATGTTGTCGTCACCAAGAGCTACTATCGATACGTTGACAGGAA 300
Db 241 GAATACATCCAACTAATGTTGTCGTCACCAAGAGCTACTATCGATACGTTGACAGGAA 300
Qy 301 CAATCATGCGGACCAACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
Db 301 CAATCATGCGGACCAACCAAAATGCAACAGTTTCGGTATCTCAAACTATTGCCAAATTTAC 360
Qy 361 CCACCAAAATGCAAAACAAAATTCGTGAAGCTTTGGCTCAAAACCCACAGCGTATTGCCGTC 420
Db 361 CCACCAAAATGTAACAAAATTTGAAGAAGCTTTGGCTCAAAACCCACAGCGCTATTGCCGTC 420
Qy 421 ATTATTGCGCATCAAGATTTAGACGCATTCGGTCAATTTATGATGCGCAAGATCAATCAA 480
Db 421 ATTATTGCGCATCAAGATTTAGACGCATTCGGTCAATTTATGATGCGCAAGATCAATCAA 480
Qy 481 CGCGAATGTTTACCAACCAAACTATCAGCTGTCACAACTTTGTTTACAGTAACGCA 540
Db 481 GAAGATAATGTTTACCAACCAACTATCAGCTGTCACAACTTTGTTTACAGTAACGCA 540
Qy 541 CAAAGTGTGATTTATGATGCGTAACAGATTTGGGATACCAATTTGGGGTGATTAATGGT 600
Db 541 CAAAGTGTGATTTATGATGCGTAACAGATTTGGGATACCAATTTGGGGTGATTAATGGT 600
Qy 601 TAGCGTTATTGCTGCCAACATCGATTTTGGATGATGATTAAGAAATATCCATATGTTGTC 660
```

```
Db      601 TACGGTATTGTTGCTGCACATCGATTGATGATGATTGAGAAATATCCATATGTTGC 660
Qy      661 ATTCTC 666
Db      661 ATTCTC 666
```

## RESULT 15

```
US-10-297-563-8/c
; Sequence 8, Application US/10297563
; Publication No. US20030138441A1
; GENERAL INFORMATION:
; APPLICANT: Alex Bollen
; APPLICANT: Paul Jacobs
; APPLICANT: Alain Jaquet
; APPLICANT: Marc Georges Francis Massier
; TITLE OF INVENTION: Codon Optimised Recombinant
; TITLE OF INVENTION: Dermaphagoides Allergens
; FILE REFERENCE: B45219
; CURRENT APPLICATION NUMBER: US/10/297,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/EP01/06483
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0014288.5
; PRIOR FILING DATE: 2000-06-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon Optimised Dermaphagoides gene
US-10-297-563-8
```

```
Query Match      11.2%; Score 74.4; DB 15; Length 100;
Best Local Similarity 84.0%; Pred. No. 3.2e-12;
Matches 84; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy      194 GTGCTTCCCAACACCGTGTGTCATGTCATACCATTCACGTCGTTGTTGAATACATCCAAAC 253
Db      100 GCGCTCCCAACACCGGATGTCATGGGATACGATCCAGAGGTATCGATATCATCCAGC 41

Qy      254 ATAATGTGTGTCGTCACCAAGAAAGTACTATCGATACGTTGC 293
Db      40 ATAATGCGTGTGTCAGGAAAGCTATTACCGATACGTAGC 1
```

Search completed: October 29, 2004, 06:26:11  
Job time : 6111.36 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:57 ; Search time 78.7171 Seconds  
(without alignments)  
187.032 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSIGNAPAEIDLRQMR.....YFAANIDLMIEEYPYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pbp:\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pbp:\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pbp:\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pbp:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pbp:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pbp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1206	100.0	245	1	US-07-945-288-2
2	1206	100.0	245	1	US-08-462-831-2
3	1206	100.0	245	1	US-08-461-809-2
4	1206	100.0	245	1	US-08-461-441-2
5	1206	100.0	245	2	US-08-482-142-2
6	1206	100.0	245	2	US-08-478-572-2
7	1206	100.0	245	3	US-08-484-296-2
8	1206	100.0	245	5	PCT-US93-08518-2
9	1206	100.0	320	1	US-07-945-288-10
10	1206	100.0	320	1	US-08-462-831-10
11	1206	100.0	320	1	US-08-461-809-10
12	1206	100.0	320	1	US-08-461-441-10
13	1206	100.0	320	5	PCT-US93-08518-10
14	1177	97.6	222	1	US-07-945-288-11
15	1177	97.6	222	1	US-08-462-831-11
16	1177	97.6	222	1	US-08-461-809-11
17	1177	97.6	222	1	US-08-461-441-11
18	1177	97.6	222	5	PCT-US93-08518-11
19	1144	94.9	245	3	US-08-460-040-2
20	997.5	82.7	321	1	US-07-945-288-6
21	997.5	82.7	321	1	US-08-462-831-6
22	997.5	82.7	321	1	US-08-461-809-6
23	997.5	82.7	321	1	US-08-461-441-6
24	997.5	82.7	321	2	US-08-482-142-6
25	997.5	82.7	321	2	US-08-478-572-6
26	997.5	82.7	321	3	US-08-484-296-6
27	997.5	82.7	321	5	PCT-US93-08518-6

28 372.5 30.9 181 2 US-08-482-142-195 Sequence 195, App  
29 372.5 30.9 181 2 US-08-478-572-195 Sequence 195, App  
30 372.5 30.9 181 3 US-08-484-296-195 Sequence 195, App  
31 354.5 29.4 181 2 US-08-482-142-197 Sequence 197, App  
32 354.5 29.4 181 2 US-08-478-572-197 Sequence 197, App  
33 354.5 29.4 181 3 US-08-484-296-197 Sequence 197, App  
34 323.5 26.8 329 1 US-08-330-121B-2 Sequence 2, Appli  
35 323.5 26.8 329 2 US-08-852-807-20 Sequence 20, Appli  
36 323.5 26.8 329 3 US-08-964-308-4 Sequence 4, Appli  
37 323.5 26.8 329 3 US-08-964-313-4 Sequence 4, Appli  
38 323.5 26.8 329 3 US-08-860-256A-1 Sequence 1, Appli  
39 323.5 26.8 329 3 US-09-069-138-4 Sequence 4, Appli  
40 323.5 26.8 329 3 US-08-915-095A-2 Sequence 2, Appli  
41 323.5 26.8 329 3 US-08-798-096-2 Sequence 2, Appli  
42 323.5 26.8 329 4 US-08-684-932A-36 Sequence 36, Appli  
43 323.5 26.8 329 4 US-08-798-095A-2 Sequence 2, Appli  
44 323.5 26.8 329 4 US-09-953-956-2 Sequence 2, Appli  
45 323.5 26.8 329 4 US-08-553-125A-2 Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-07-945-288-2  
; Sequence 2, Application US/07945288  
; Patent No. 5433948  
; GENERAL INFORMATION:  
; APPLICANT: Thomas, Wayne R.  
; APPLICANT: Chua, Kaw-Yan  
; TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
; TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/945,288  
FILING DATE: 19920910  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: P36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 245 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-945-288-2

Query Match 100.0%; Score 1206; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.3e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 Db 24 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 143  
 QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 180  
 Db 144 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 203  
 QY 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVVIL 222  
 Db 204 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVVIL 245

## RESULT 2

US-08-462-831-2  
 ; Sequence 2, Application US/08462831  
 ; Patent No. 5552142  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
 ; TITLE OF INVENTION: DERMATOPHAGOIDES  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/462,831  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/945,288  
 ; FILING DATE: 10 SEPTEMBER 1992  
 ; APPLICATION NUMBER: US 580,655  
 ; FILING DATE: 11 SEPTEMBER 1990  
 ; APPLICATION NUMBER: US 458,642  
 ; FILING DATE: 13 FEBRUARY 1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MANDRAGOURAS, AMY E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 245 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-462-831-2

Query Match 100.0%; Score 1206; DB 1; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-129;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 Db 24 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120

Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 143  
 QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 180  
 Db 144 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 203  
 QY 181 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVVIL 222  
 Db 204 QGVYWIVRNSWDTNWDNGYGYFAANIDLMIMEEYPYVVIL 245

## RESULT 3

US-08-461-809-2  
 ; Sequence 2, Application US/08461809  
 ; Patent No. 570202  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
 ; TITLE OF INVENTION: DERMATOPHAGOIDES  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/461,809  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/945,288  
 ; FILING DATE: 10 SEPTEMBER 1992  
 ; APPLICATION NUMBER: US 580,655  
 ; FILING DATE: 11 SEPTEMBER 1990  
 ; APPLICATION NUMBER: US 458,642  
 ; FILING DATE: 13 FEBRUARY 1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MANDRAGOURAS, AMY E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 245 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-461-809-2

Query Match 100.0%; Score 1206; DB 1; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-129;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 Db 24 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 143  
 QY 121 PPANKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 180

Db 144 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 203

QY 181 QGVYWIVRNSWDTNWGDNGYGFYFAANIDLMIMEEYPYVIL 222

Db 204 QGVYWIVRNSWDTNWGDNGYGFYFAANIDLMIMEEYPYVIL 245

RESULT 4

US-08-461-441-2

; Sequence 2, Application US/08461441

; Patent No. 5773002

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 STATE STREET, SUITE 510

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/461,441

; FILING DATE: 13 FEBRUARY 1990

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/945,288

; FILING DATE: 10 SEPTEMBER 1992

; APPLICATION NUMBER: US 580,655

; FILING DATE: 11 SEPTEMBER 1990

; APPLICATION NUMBER: US 458,642

; FILING DATE: 13 FEBRUARY 1990

; ATTORNEY/AGENT INFORMATION:

; NAME: MANDRAGOURAS, AMY E.

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 245 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-461-441-2

Query Match 100.0%; Score 1206; DB 1; Length 245;

Best Local Similarity 100.0%; Pred. No. 1.3e-129;

Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

Db 24 TNACSGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNACRFGISNYCQIY 120

Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNACRFGISNYCQIY 143

QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180

Db 144 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 203

QY 181 QGVYWIVRNSWDTNWGDNGYGFYFAANIDLMIMEEYPYVIL 222

Db 204 QGVYWIVRNSWDTNWGDNGYGFYFAANIDLMIMEEYPYVIL 245

RESULT 5

US-08-482-142-2

; Sequence 2, Application US/08482142

; Patent No. 5820862

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: MA

; COUNTRY: USA

; ZIP: 02154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/482,142

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/445,307

; FILING DATE: 07 June 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG, ANNE I.

; REGISTRATION NUMBER: 32,976

; REFERENCE/DOCKET NUMBER: 017.6US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 466-6000

; TELEFAX: (617) 466-6040

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 245 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-482-142-2

Query Match 100.0%; Score 1206; DB 2; Length 245;

Best Local Similarity 100.0%; Pred. No. 1.3e-129;

Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

Db 24 TNACSGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNACRFGISNYCQIY 120

Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNACRFGISNYCQIY 143

QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180

Db 144 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 203

QY 181 QGVYWIVRNSWDTNWGDNGYGFYFAANIDLMIMEEYPYVIL 222

Db 204 QGVYWIVRNSWDTNWGDNGYGFYFAANIDLMIMEEYPYVIL 245

RESULT 6  
US-08-478-572-2

; Sequence 2, Application US/08478572  
; Patent No. 5968526

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)

; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: MA

; COUNTRY: USA

; ZIP: 02154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/478,572

; FILING DATE: 07-June-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/445,307

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG, ANNE I.

; REGISTRATION NUMBER: 32,976

; REFERENCE/DOCKET NUMBER: 017.6US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 466-6000

; TELEFAX: (617) 466-6040

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 245 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-478-572-2

Query Match 100.0%; Score 1206; DB 2; Length 245;

Best Local Similarity 100.0%; Pred. No. 1.3e-129;

Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSIGNAPAEIDLEQMTVTPIRMQCGCGSCNAPFGVAATESAYLAHRNQSLLDAEQ 60

DB 24 TNACSIGNAPAEIDLEQMTVTPIRMQCGCGSCNAPFGVAATESAYLAHRNQSLLDAEQ 83

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120

DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 143

QY 121 PPNAKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNHVNIVGYSNA 180

DB 144 PPNAKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNHVNIVGYSNA 203

QY 181 QGVYDWIVRNSWDTNWGDNCGYGYFAANIDLMWIEEYPYVIL 222

DB 204 QGVYDWIVRNSWDTNWGDNCGYGYFAANIDLMWIEEYPYVIL 245

RESULT 7

US-08-484-296-2

; Sequence 2, Application US/08484296

; Patent No. 6268491

; GENERAL INFORMATION:

; APPLICANT: Garman, Richard

; APPLICANT: Greenstein, Julia

; APPLICANT: Kuo, Mei-chang

; APPLICANT: Rogers, Bruce

; APPLICANT: Franzen, Henry

; APPLICANT: Chen, Xian

; APPLICANT: Evans, Sean

; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS

; FROM DERMATOPHAGOIDES (HOUSE DUST MITE)

; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: MA

; COUNTRY: USA

; ZIP: 02154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/484,296

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/445,307

; FILING DATE: 07 June 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: CRAIG, ANNE I.

; REGISTRATION NUMBER: 32,976

; REFERENCE/DOCKET NUMBER: 017.6US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 466-6000

; TELEFAX: (617) 466-6040

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 245 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-484-296-2

Query Match 100.0%; Score 1206; DB 3; Length 245;

Best Local Similarity 100.0%; Pred. No. 1.3e-129;

Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSIGNAPAEIDLEQMTVTPIRMQCGCGSCNAPFGVAATESAYLAHRNQSLLDAEQ 60

DB 24 TNACSIGNAPAEIDLEQMTVTPIRMQCGCGSCNAPFGVAATESAYLAHRNQSLLDAEQ 83

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 120

DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAORFGISNYCQIY 143

QY 121 PPNAKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNHVNIVGYSNA 180

DB 144 PPNAKIREALAQTHSAIAVIIIGIKDLDAFRHYDGRITIIQRDNGYQPNHVNIVGYSNA 203

QY 181 QGVYDWIVRNSWDTNWGDNCGYGYFAANIDLMWIEEYPYVIL 222

DB 204 QGVYDWIVRNSWDTNWGDNCGYGYFAANIDLMWIEEYPYVIL 245

RESULT 8

PCT-US93-08518-2

; Sequence 2, Application PC/TUS9308518

; GENERAL INFORMATION:

APPLICANT: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM: FLOPPY disk  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08518  
FILING DATE: 10 SEPTEMBER 1992  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 245 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-08518-2

Query Match 100.0%; Score 1206; DB 5; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.3e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACSGNAPAEIDLQRMRTVPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDAEQ 60  
Db 24 TNACSGNAPAEIDLQRMRTVPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDAEQ 83  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHGVVQESYRYVAREQSCRRPNQRFISNYCQIY 120  
Db 84 ELVDCASQHGCHGDTIPRGIEYIQHGVVQESYRYVAREQSCRRPNQRFISNYCQIY 143  
Qy 121 PPANKIREALAOQTHSAIAVIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
Db 144 PPANKIREALAOQTHSAIAVIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 203  
Qy 181 QGVYWIWRNSWDTNWGDNGYGYFAANIDLMMEIEYPPYVIL 222  
Db 204 QGVYWIWRNSWDTNWGDNGYGYFAANIDLMMEIEYPPYVIL 245

RESULT 9  
US-07-945-288-10  
Sequence 10, Application US/07945288  
Patent No. 5433948  
GENERAL INFORMATION:  
APPLICANT: Thomas, Wayne R.  
APPLICANT: Chua, Kaw-Yan  
TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/945,288  
FILING DATE: 19920910  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-945-288-10

Query Match 100.0%; Score 1206; DB 1; Length 320;  
Best Local Similarity 100.0%; Pred. No. 1.9e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNACSGNAPAEIDLQRMRTVPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDAEQ 60  
Db 99 TNACSGNAPAEIDLQRMRTVPIRMGGCGSCWAFSGVAATESAYLAHRNQSILDAEQ 158  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHGVVQESYRYVAREQSCRRPNQRFISNYCQIY 120  
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHGVVQESYRYVAREQSCRRPNQRFISNYCQIY 218  
Qy 121 PPANKIREALAOQTHSAIAVIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
Db 219 PPANKIREALAOQTHSAIAVIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 278  
Qy 181 QGVYWIWRNSWDTNWGDNGYGYFAANIDLMMEIEYPPYVIL 222  
Db 279 QGVYWIWRNSWDTNWGDNGYGYFAANIDLMMEIEYPPYVIL 320

RESULT 10  
US-08-462-831-10  
Sequence 10, Application US/08462831  
Patent No. 5552142  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,831
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-831-10

```

```

Query Match 100.0%; Score 1206; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.9e-129; Indels 0; Gaps 0;
Matches 222; Conservative 0; Mismatches 0;

QY 1 TNACISNGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
DB 99 TNACISNGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 158

QY 61 ELVDCASQHCCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAORFGISNYCOIY 120
DB 159 ELVDCASQHCCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAORFGISNYCOIY 218

QY 121 PPNAKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
DB 219 PPNAKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278

QY 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMMEIEEYPYVIL 222
DB 279 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMMEIEEYPYVIL 320

```

```

RESULT 11
US-08-461-809-10
; Sequence 10, Application US/08461809
; Patent No. 5770202
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,809
; FILING DATE:
; CLASSIFICATION: 424

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-809-10

Query Match 100.0%; Score 1206; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.9e-129; Indels 0; Gaps 0;
Matches 222; Conservative 0; Mismatches 0;

QY 1 TNACISNGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
DB 99 TNACISNGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 158

QY 61 ELVDCASQHCCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAORFGISNYCOIY 120
DB 159 ELVDCASQHCCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAORFGISNYCOIY 218

QY 121 PPNAKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
DB 219 PPNAKIREALAOQTHSAIAVIIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278

QY 181 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMMEIEEYPYVIL 222
DB 279 QGVVDYWIVRNSWDTNWDNGYGYFAANIDLMMEIEEYPYVIL 320

```

```

RESULT 12
US-08-461-441-10
; Sequence 10, Application US/08461441
; Patent No. 5773002
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; TITLE OF INVENTION: DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,441
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990

```

APPLICATION NUMBER: US 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
US-08-461-441-10

Query Match 100.0%; Score 1206; DB 1; Length 320;  
Best Local Similarity 100.0%; Pred. No. 1.9e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATSAYLAAHRNQSILDLAEQ 60  
DB 99 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATSAYLAAHRNQSILDLAEQ 158  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
DB 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGISNYCQIY 218  
QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
DB 219 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 278  
QY 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMMEIEYPYVIL 222  
DB 279 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMMEIEYPYVIL 320

RESULT 13  
PCT-US93-08518-10  
Sequence 10, Application PC/TUS9308518  
GENERAL INFORMATION:  
APPLICANT: CHUA, KAW-YAN  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08518  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/945,288  
FILING DATE: 10 SEPTEMBER 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-08518-10

Query Match 100.0%; Score 1206; DB 5; Length 320;  
Best Local Similarity 100.0%; Pred. No. 1.9e-129;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATSAYLAAHRNQSILDLAEQ 60  
DB 99 TNACSSINGNAPAEIDLQRMRTVTPIRMQGGCGSCWAFSGVAATSAYLAAHRNQSILDLAEQ 158  
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
DB 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVVQESYRYVAREQSCRRPNAQRFGISNYCQIY 218  
QY 121 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
DB 219 PPNAKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 278  
QY 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMMEIEYPYVIL 222  
DB 279 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMMEIEYPYVIL 320

RESULT 14  
US-07-945-288-11  
Sequence 11, Application US/07945288  
Patent No. 5433948  
GENERAL INFORMATION:  
APPLICANT: THOMAS, WAYNE R.  
APPLICANT: CHUA, KAW-YAN  
TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/945,288  
FILING DATE: 19920910  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: P36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:

NAME/KEY: misc feature  
LOCATION: 50  
OTHER INFORMATION: /label=Xaa is His or Tyr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 81  
OTHER INFORMATION: /label=Xaa is Glu or Lys  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 124  
OTHER INFORMATION: /label=Xaa is Ala or Val  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 136  
OTHER INFORMATION: /label=Xaa is Ser or Thr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 215  
OTHER INFORMATION: /label=Xaa is Glu or Gln  
US-07-945-288-11

Query Match 97.6%; Score 1177; DB 1; Length 222;  
Best Local Similarity 97.7%; Pred. No. 2.3e-126;  
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 TNACSSINGNAPAEIDLQRMETVTPIRMQCGSCSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Db 1 TNACSSINGNAPAEIDLQRMETVTPIRMQCGSCSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIXYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
Qy 121 PPNANKIREALAOATHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHANNVIGYSNA 180  
Db 121 PPNANKIREALAOATHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHANNVIGYSNA 180  
Qy 181 QGVDDYWIVRNSWDTNWGDNCGYGFANIDLMITEEYPPYVIL 222  
Db 181 QGVDDYWIVRNSWDTNWGDNCGYGFANIDLMITEEYPPYVIL 222

RESULT 15  
US-08-462-831-11  
Sequence 11, Application US/08462831  
Patent No. 5552142  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,831  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/945,288  
FILING DATE: 10 SEPTEMBER 1992  
APPLICATION NUMBER: US 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: US 458,642

FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 35,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 50  
OTHER INFORMATION: /label=Xaa is His or Tyr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 81  
OTHER INFORMATION: /label=Xaa is Glu or Lys  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 124  
OTHER INFORMATION: /label=Xaa is Ala or Val  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 136  
OTHER INFORMATION: /label=Xaa is Ser or Thr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 215  
OTHER INFORMATION: /label=Xaa is Glu or Gln  
US-08-462-831-11

Query Match 97.6%; Score 1177; DB 1; Length 222;  
Best Local Similarity 97.7%; Pred. No. 2.3e-126;  
Matches 217; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 TNACSSINGNAPAEIDLQRMETVTPIRMQCGSCSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Db 1 TNACSSINGNAPAEIDLQRMETVTPIRMQCGSCSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
Db 61 ELVDCASQHGCHGDTIPRGIXYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
Qy 121 PPNANKIREALAOATHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHANNVIGYSNA 180  
Db 121 PPNANKIREALAOATHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHANNVIGYSNA 180  
Qy 181 QGVDDYWIVRNSWDTNWGDNCGYGFANIDLMITEEYPPYVIL 222  
Db 181 QGVDDYWIVRNSWDTNWGDNCGYGFANIDLMITEEYPPYVIL 222

Search completed: October 27, 2004, 17:53:00  
Job time : 79.7171 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:12 ; Search time 59.259 Seconds  
(without alignments)  
360.454 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSSINGNAPAEIDLQRM.....YFAANIDLMMIEYPYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1144	94.9	245	JQ0337	allergen Der p 1 -
2	997.5	82.7	319	A81500	allergen Der f 1 p
3	982.5	81.5	211	S21864	probable cysteine
4	389.5	32.3	94	S03380	major fecal allerg
5	333.5	27.7	358	JC7787	carrot seed cystei
6	332	27.5	348	F84672	probable cysteine
7	327	27.1	348	JN0633	caricain (EC 3.4.2
8	327	27.1	367	JN0634	caricain (EC 3.4.2
9	326.5	27.1	326	T09259	cathepsin L-like p
10	323.5	26.8	329	JC2476	cathepsin K (EC 3.
11	322.5	26.7	374	T03941	cysteine proteinas
12	322	26.7	346	C86413	cysteine proteinas
13	322	26.7	362	S12581	cysteine proteinas
14	321	26.6	368	S47312	cysteine proteinas
15	321	26.6	466	T06416	cysteine proteinas
16	317.5	26.3	329	A49868	cathepsin K (EC 3.
17	315	26.1	215	A59428	ervatamin B (EC 3.
18	313	26.0	458	1 KHRZOA	oryzain (EC 3.4.22
19	312.5	25.9	342	S71773	cysteine proteinas
20	311.5	25.8	312	A23705	cysteine proteinas
21	311.5	25.8	329	S74277	cathepsin K (EC 3.
22	311	25.8	221	A59041	cysteine proteinas
23	311	25.8	222	S13649	cysteine proteinas
24	310.5	25.7	380	1 TAGB	actinidin (EC 3.4
25	310	25.7	454	2 JC4848	cysteine proteinas
26	310	25.7	471	1 KHRZOB	oryzain (EC 3.4.22
27	309.5	25.7	326	S43991	cathepsin L-like p
28	308.5	25.6	323	2 S19650	cysteine proteinas
29	308	25.5	365	2 T06205	probable cysteine

30 308 25.5 367 2 T12382 cysteine proteinas  
31 307.5 25.5 345 2 T07839 ananain (EC 3.4.22  
32 307 25.5 352 2 T09760 chymopapain (EC 3.  
33 307 25.5 385 2 T06208 cysteine proteinas  
34 307 25.5 367 2 T06529 drought-inducible  
35 307 25.5 462 2 JN0719 cysteine proteinas  
36 306 25.4 362 2 S22502 cysteine proteinas  
37 305.5 25.3 356 2 T30426 actinidin-like pro  
38 304.5 25.2 302 2 S02728 cathepsin (EC 3.4  
39 304.5 25.2 325 2 S49451 cysteine proteinas  
40 304.5 25.2 361 2 T06708 cysteine proteinas  
41 304 25.2 313 2 S47433 cathepsin L (EC 3.  
42 303.5 25.2 328 2 JQ1121 cysteine proteinas  
43 303 25.1 348 2 T09798 glycy l endopeptida  
44 303 25.1 349 2 T06207 cysteine proteinas  
45 303 25.1 364 2 T46630 cysteine proteinas

## ALIGNMENTS

### RESULT\_1

JQ0337  
allergen Der p 1 - house-dust mite (Dermatophagoides pteronyssinus) (fragment)  
C;Species: Dermatophagoides pteronyssinus  
C;Date: 07-Sep-1990 #sequence revision 07-Sep-1990 #text\_change 09-Jul-2004  
C;Accession: JQ0337; A27582; A31657; C27634  
R;Chua, K.Y.; Stewart, G.A.; Thomas, W.R.; Simpson, R.J.; Plozza, T.M.; J  
J. Exp. Med. 167, 175-182, 1988  
A;Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1:  
A;Reference number: JQ0337; MUID:88089411; PMID:3335830  
A;Accession: JQ0337  
A;Molecule type: mRNA  
A;Residues: 1-245 <CHU>  
A;Cross-references: UNIPROT:P08176  
R;Thomas, W.R.; Stewart, G.A.; Simpson, R.J.; Chua, K.Y.; Plozza, T.M.; Dilworth, R.J.; J  
Int. Arch. Allergy Appl. Immunol. 85, 127-129, 1988  
A;Title: Cloning and expression of DNA coding for the major house dust mite allergen Der  
A;Reference number: A27582; MUID:88114080; PMID:3276629  
A;Accession: A27582  
A;Molecule type: mRNA  
A;Residues: 6-101 <THO>  
A;Cross-references: GB:M24794; NID:G387591; PIDN:AAA28296.1; PID:G387592  
R;Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.  
Protein Seq. Data Anal. 2, 17-21, 1989  
A;Title: Structural studies on the allergen Der p1 from the house dust mite Dermatophagoid  
A;Reference number: A31657; MUID:89098855; PMID:2911558  
A;Accession: A31657  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 24-55, 'XX', 59-64;102-118;134-149;185-192, 'X', 194-200, 'X', 202 <SIM>  
R;Lind, P.; Hansen, O.C.; Horn, N.  
J. Immunol. 140, 4256-4262, 1988  
A;Title: The binding of mouse hyalidoma and human IgE antibodies to the major fecal aller:  
d by solid-phase inhibition assays with radiolabeled antigen.  
A;Reference number: A92819; MUID:88229138; PMID:3372999

A;Accession: C27634  
A;Molecule type: protein  
A;Residues: 24-52 <LIN>  
C;Superfamily: papain  
C;Keywords: Glycoprotein  
F;75/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 94.9%; Score 1144; DB 2; Length 245;  
Best Local Similarity 95.5%; Pred. No. 1.5e-95;  
Matches 212; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 TNACSSINGNAPAEIDLQRMRTVTPIRMQCGGSCWAFSGVAATESAYLAHRNQSLEAEQ 60

Db 24 TNACSSINGNAPAEIDLQRMRTVTPIRMQCGGSCWAFSGVAATESAYLAHRNQSLEAEQ 83

Qy 61 ELVDCASQHCCHDITPRGIEVTHQNGVVQESYRVVAREQSCRRPNARFQISNYCQY 120

Db 84 ELVDCAHQCHGDTIPRGIEYIQHNGVQESYRYVAREQCRPNQAFGLSNYCOI 143  
 Qy 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSNA 180  
 Db 144 PPNANKIREALQPYRCRHYWTIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSNA 203  
 Qy 181 QGVDDYIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVIL 222  
 Db 204 QGVDDYIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 2  
 A61500  
 allergen Der f I precursor - house-dust mite (Dermatophagoides farinae)  
 C:Species: Dermatophagoides farinae  
 C>Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 26-Aug-1999  
 C:Accession: A61500  
 R:Dilworth, R.J.; Chua, K.Y.; Thomas, W.R.  
 Clin. Exp. Allergy 21, 25-32, 1991  
 A:Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I.  
 A:Reference number: A61500; MUID:91215493; PMID:2021874  
 A:Accession: A61500  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-319 <DIL>  
 C:Superfamily: papain

Query Match 82.7%; Score 997.5; DB 2; Length 319;  
 Best Local Similarity 81.2%; Pred. No. 3.2e-82;  
 Matches 181; Conservative 18; Mismatches 23; Indels 1; Gaps 1;  
 Qy 1 TNACINSVNPSELDLSRLTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSILDLAE 59  
 Db 97 TSACRINSVNPSELDLSRLTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSILDLAE 156  
 Qy 60 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQCRPNQAFGLSNYCOI 119  
 Db 157 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQCRPNQAFGLSNYCOI 216  
 Qy 120 YPPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSN 179  
 Db 217 YPPDKQIRALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSN 276  
 Qy 180 AQGVDDYIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVIL 222  
 Db 277 TQGVDDYIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVIL 319

RESULT 3  
 S21864  
 probable cysteine proteinase (BC 3.4.22.-) - Euroglyphus maynei  
 N:Alternate names: allergen Eur m I  
 C:Species: Euroglyphus maynei  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C:Accession: S21864  
 R:Kent, N.A.; Hill, M.; Keen, J.N.; Holland, P.W.H.; Hart, B.  
 submitted to the EMBL Data Library, June 1991  
 A:Reference number: S21864  
 A:Accession: S21864  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-211 <KEN>  
 C:Cross-references: UNIPROT:P25780; EMBL:X60073  
 C:Genetics:  
 A:Introns: 100/3; 155/2  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase

Query Match 81.5%; Score 982.5; DB 2; Length 211;  
 Best Local Similarity 83.9%; Pred. No. 4.3e-81;  
 Matches 177; Conservative 17; Mismatches 16; Indels 1; Gaps 1;

Qy 1 TNACINSVNPSELDLSRLTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSILDLAE 59  
 Db 1 TYACINSVNPSELDLSRLTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSILDLAE 60  
 Qy 60 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQCRPNQAFGLSNYCOI 119  
 Db 61 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQCRPNQAFGLSNYCOI 120  
 Qy 120 YPPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSN 179  
 Db 121 SPDSNKIRALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSN 180  
 Qy 180 AQGVDDYIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVIL 222  
 Db 181 TQGVDDYIVRNSWDTNWGDNGYGYFAANIDLMIMEEYPYVIL 319

RESULT 4  
 S03380  
 major fecal allergen Der p I - house-dust mite (Dermatophagoides pteronyssinus) (fragment)  
 N:Alternate names: allergen Der p I  
 C:Species: Dermatophagoides pteronyssinus  
 C>Date: 05-Mar-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
 C:Accession: S03380  
 R:Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.  
 Protein Seq. Data Anal. 2, 17-21, 1989  
 A:Title: Structural studies on the allergen Der p I from the house dust mite Dermatophagoides pteronyssinus  
 A:Reference number: A31657; MUID:89098855; PMID:2911558  
 A:Accession: S03380  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-28; 29-43; 44-60; 61-76; 77-94 <SIM>  
 A:Cross-references: UNIPROT:Q7M431  
 C:Superfamily: papain

Query Match 32.3%; Score 389.5; DB 2; Length 94;  
 Best Local Similarity 48.6%; Pred. No. 4.2e-28;  
 Matches 88; Conservative 0; Mismatches 4; Indels 89; Gaps 4;  
 Qy 1 TNACINSVNPSELDLSRLTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSILDLAE 58  
 Db 1 TNACINSVNPSELDLSRLTVPTRMQGCGSCWAFSGVAATESAYLAHRNQSILDLAE 43  
 Qy 59 QELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQCRPNQAFGLSNYCOI 118  
 Db 44 -----GIEYIQHNGVQESY-----RGISNYCO 68  
 Qy 119 YPPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIGYSN 178  
 Db 69 IYPENANK-----DNGYQPNYHVNIGYSN 93

RESULT 5  
 JC7787  
 carrot seed cysteine proteinase (EC 3.4.-.-), CSCP - carrot  
 C:Species: Daucus carota (carrot)  
 C>Date: 02-Apr-2002 #sequence\_revision 02-Apr-2002 #text\_change 09-Jul-2004  
 C:Accession: JC7787  
 R:Sakuta, C.; Oda, A.; Konishi, M.; Yamakawa, S.; Kamada, H.; Satoh, S.  
 Biosci. Biotechnol. Biochem. 65, 2243-2248, 2001  
 A:Title: Cysteine proteinase gene expression in the endosperm of germinating carrot seeds  
 A:Reference number: JC7787  
 A:Contents: Seeds  
 A:Accession: JC7787  
 A:Molecule type: mRNA  
 A:Residues: 1-358 <SAK>  
 A:Cross-references: UNIPROT:Q948S1; DBJ:AB057371; DDBJ:AB057372  
 C:Comment: This protein is concerned with programmed cell death, and is endosperm-specific  
 C:Genetics:

C;Species: Carica papaya (papaya)  
C;Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: JN0633, S01135, S11748; A43026, S35728  
R;Revell, D.F.; Cummings, N.J.; Baker, K.C.; Collins, M.E.; Taylor, M.A.J.; Sumner, I.G.;  
Gene 127, 221-225, 1993  
A;Title: Nucleotide sequence and expression in Escherichia coli of cDNAs encoding papaya I  
A;Reference number: JN0633; MUID:93273235; PMID:7684720  
A;Accession: JN0633  
A;Molecule type: mRNA  
A;Residues: 1-348 <REV>  
A;Cross-references: UNIPROT:P10056; EMBL:X66060; NID:gl18097; PIDN:CAA46862.1; PID:gl18098  
A;Experimental source: fruit and leaf  
R;Dubois, T.; Kleinschmidt, T.; Schnek, A.G.; Looze, Y.; Braunitzer, G.  
Biochem. J. 261, 189-193, 1989  
A;Title: The thiol proteinases from the latex of Carica papaya L. II. The primary structure  
A;Reference number: S01135; MUID:89105320; PMID:3063283  
A;Accession: S01135  
A;Molecule type: protein  
A;Residues: 133-348 <DUB>  
R;Collins, M.E.; Revell, D.F.; Sumner, I.G.; Pickersgill, R.W.; Goodenough, P.W.  
submitted to the EMBL Data Library, February 1990  
A;Reference number: S11748  
A;Accession: S11748  
A;Molecule type: mRNA  
A;Residues: 237-348 <COL>  
A;Cross-references: EMBL:X51899; NID:gl18095; PIDN:CAA36180.1; PID:gl360713  
C;Genetics:  
A;Gene: Pp-omega  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;1-14/Domain; signal sequence #status predicted <SIG>  
F;15-132/Domain; propeptide #status predicted <PRO>  
F;133-348/Product; caricain #status experimental <MAT>  
F;154-195/188-227, 285-336/Disulfide bonds; #status predicted  
F;157,291,311/Active site: Cys, His, Asn #status predicted

Query Match 27.1%; Score 327; DB 2; Length 348;  
Best Local Similarity 37.2%; Pred. No. 8.4e-22;  
Matches 74; Conservative 34; Mismatches 77; Indels 14; Gaps 7;

QY 9 NAPAELDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDLAEQELVDCASQ 68  
Db 132 NLPEVDWRKKGAVTPVRHQGGCGSWAFSAVATGINKIRTKGLVELSEQLVDCERR 191

QY 69 -HCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFQ-----ISNYCOIYPP 122  
Db 192 SHGCKGGYPYALEYVAKNGIHLRSKYPYAKQGTCT--AKQVGGPIVKTSGVGRVQFN 248

QY 123 NANKIREALQAQTHSAIAVIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYVNAQ 182  
Db 249 NEGNLLNATK--QPVSVVVEKGR-PPFQLYKG-GIFEGPCGTKVD-HAVTAVGYGKSG 303

QY 183 VDIWIVNSWDTNWGDNGY 201  
Db 304 KGYILIKNSWGTAWGKGY 322

RESULT 8  
JN0634  
caricain (EC 3.4.22.30) II precursor - papaya  
N;Alternate names: papaya proteinase omega II  
C;Species: Carica papaya (papaya)  
C;Date: 03-Feb-1994 #sequence revision 03-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: JN0634; A43027; S31823  
R;Revell, D.F.; Cummings, N.J.; Baker, K.C.; Collins, M.E.; Taylor, M.A.J.; Sumner, I.G.;  
Gene 127, 221-225, 1993  
A;Title: Nucleotide sequence and expression in Escherichia coli of cDNAs encoding papaya I  
A;Reference number: JN0633; MUID:93273235; PMID:7684720  
A;Accession: JN0634  
A;Molecule type: mRNA  
A;Residues: 1-367 <REV>  
A;Cross-references: UNIPROT:Q42673; EMBL:X69877; NID:g22660; PIDN:CAA49504.1; PID:g22661  
A;Experimental source: fruit and leaf

A;Gene: cscp  
C;Superfamily: papain  
C;Keywords: cysteine proteinase; hydrolase  
F;151,286,307/Active site: Cys, His, Asn #status predicted

Query Match 27.7%; Score 333.5; DB 2; Length 358;  
Best Local Similarity 34.8%; Pred. No. 2.3e-22;  
Matches 77; Conservative 36; Mismatches 87; Indels 21; Gaps 7;

QY 11 PAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDLAEQELVDCASQ- 69  
Db 128 FSSVDRKHGAVTPFKNGRCGWSAFSAIVGVEGINKINKTNQLVLSUSEQELVDCESDQ 187

QY 70 GCHGDTIPRGIEYIQHNGVQESYRYVAREQSC--RRPNAQRFGLSNYCOIYPPNANK 126  
Db 188 GCGGLMENALEFKRSGGVTVRYPYARDERCDAKGNAPVVKIDGHENVPENNEYA 247

QY 127 IREALQAQTHSAIAVIGIKDLDAFRH--YDGRTHQDRNGYQPNYHVNIVGYVNAQGV 183  
Db 248 LAQAVANQFVSAIDAGSDMQFYREGVYTGECGTELD-----HGVAVVGATNDGT 300

QY 184 DYWIVRNSWDTNWGDNGYVFAANID-----LMMIEEYP 217  
Db 301 KYWIVRNSWGTWGERGYIRWRDINAEIGCIAMEASYP 341

RESULT 6  
F84672  
probable cysteine proteinase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: F84672  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: F84672  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-348 <STO>  
A;Cross-references: UNIPROT:Q9ZQH7; GB:AE002093; NID:g4314384; PIDN:AAD15594.1; GSPDB:GN  
C;Genetics:  
A;Gene: At2g27420  
A;Map position: 2  
C;Superfamily: papain

Query Match 27.5%; Score 332; DB 2; Length 348;  
Best Local Similarity 36.2%; Pred. No. 3e-22;  
Matches 84; Conservative 32; Mismatches 86; Indels 30; Gaps 11;

QY 5 SINGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDLAEQELV 64  
Db 126 SDNGES---MDWRQEGAVTPVKYQRCGCGWAFSAVAAVEGITKTKGLVSLSEQLLD 182

QY 65 CASQH--GCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNA-----QRFGISNY 116  
Db 183 CDRDYNQRCGGINSKAEYIKNGITENYFYQESQQTCSSSTLSSSFRAATISGY 242

QY 117 COIYPPNANKIREALQAQTHSAIAVIGIKDLDAFRHYDGTIIQRDNGY--QPNYHVN 173  
Db 243 -ETVPMNN---EALLQAVSQQPVSVGIEGTGAAPRHSYGGVF---NGECGTDLHHAFT 294

QY 174 IVGYS-NAQGVYDIWIVRNSWDTNWGDNGYVFAANID-----LMMIEEYP 217  
Db 295 IVGYGMSSEGTKYVWVKNWGTWGTGNGYIMRIKRDVDAPQMGCLAILAFTP 346

RESULT 7  
JN0633  
caricain (EC 3.4.22.30) I precursor - papaya  
N;Alternate names: papaya proteinase omega I; peptidase A; peptidase II; proteinase A; P

```
C;Genetics:
A;Gene: Pp-omega
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase
F;1-11/Domain: signal sequence #status predicted <SIG>
F;112-132/Domain: propeptide #status predicted <PRO>
F;133-367/Product: carnicain II #status predicted <MAT>
F;154-195,188-227,265-336/Disulfide bonds: #status predicted
F;157,291,311/Active site: Cys, His, Asn #status predicted

Query Match      27.1%; Score 327; DB 2; Length 367;
Best Local Similarity 37.2%; Pred. No. 9e-22;
Matches 74; Conservative 34; Mismatches 77; Indels 14; Gaps 7;

QY 9 NAPAETDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAQELVDCASQ 68
DB 132 NLBNVDWRKKGAVTPVRQGGCGSCWAFSAVATVEGINKIRTKLVLSQELVDCERR 191
QY 69 -HCHGDTIPRGTEYIQHNQVQESYRYVAREQSCRRPNAQRFQ-----ISNYCOIYPP 122
DB 192 SHGCKGGYPYALEYVAKNGIHLSKYPYKAKQGTCTCR---AKVGGPIVKTSGVGRVQPN 248
QY 123 NANKIREALQTHSAIAVIGIKDLDAFRHYDGRITIORDNQYQPNYHVNITVGSNAQ 182
DB 249 NEGNLNLAIK--QPVSVVVEKSG-RPFQLYKG-GIFEGPCGTQVD-HAVTAVGYKSG 303
QY 183 VDYWIVRNSWDTNWGDNGY 201
DB 304 KGVILKNSMGTAWEKGY 322

RESULT 9
T09259
cathepsin L-like proteinase (EC 3.4.22.-) - liver fluke
C;Species: Fasciola hepatica (liver fluke)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T09259
R;Heusler, V.T.; Dobbelaere, D.A.E.
Mol. Biochem. Parasitol. 64, 11-23, 1994
A;Title: Cloning of a protease gene family of Fasciola hepatica by polymerase chain reaction
A;Reference number: Z16631; MUID:94359526; PMID:8078514
A;Accession: T09259
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-326 <HEU>
A;Cross-references: UNIPROT:Q24944; EMBL:Z22765; NID:G452257; PID:G452258
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase
F;132,269,289/Active site: Cys, His, Asn #status predicted

Query Match      27.1%; Score 326.5; DB 2; Length 326;
Best Local Similarity 34.4%; Pred. No. 8.7e-22;
Matches 72; Conservative 31; Mismatches 83; Indels 23; Gaps 5;

QY 11 PAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAQELVDCASQ 67
DB 109 PESIDWRDYYVTEVQDQGGCGSCWAFSTTGAVEGQFRKNERASASFSQQLVDCTRDFG 168
QY 68 QHCHGDTIPRGTEYIQHNQVQESYRYVAREQSCRRPNAQRFQ-ISNYCOIYPPNANK 126
DB 169 NYCGGGYMNAYEYLKHNGLTESYTPYQAVGPGCGRLAIYAKVGYTVVHSGDEIE 228
QY 127 IREALQTHSAIAVIGIKDLDAFR-HYDGRITIORDNQYQPN-----YHAVNITVGS 178
DB 229 LKN-----LVGTEDLPAVALDADSDPMMYQSGIYQSTCLPDLTHAVLAAGVG 277
QY 179 NAGVDVWIVRNSWDTNWGDNGYGYFAAN 207
DB 278 SODGTDWYIVKNSMGTWGDEGYIRFARN 306

cathepsin K (EC 3.4.22.-) precursor - human
N;Alternate names: cathepsin O2
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C;Accession: JC2476; S55763; S68459; I38752; S48830
R;Inoka, T.; Bilbe, G.; Ishibashi, O.; Tezuka, K.; Kumegawa, M.; Kokubo, T.
Biochem. Biophys. Res. Commun. 206, 89-96, 1995
A;Title: Molecular cloning of human cDNA for cathepsin K: Novel cysteine proteinase predicted
A;Reference number: JC2476; MUID:95118380; PMID:7818555
A;Accession: JC2476
A;Molecule type: mRNA
A;Residues: 1-329 <INA>
A;Cross-references: UNIPROT:P43235; EMBL:X82153; NID:G562756; PIDN:CAA57649.1; PID:G562756
R;Bromme, D.; Okamoto, K.
Biol. Chem. Hoppe-Seyler 376, 379-384, 1995
A;Title: Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastoma
A;Reference number: S55763; MUID:96082523; PMID:7576232
A;Accession: S55763
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-329 <BRO>
A;Cross-references: GB:S79895; NID:g1195555; PIDN:AAB35521.1; PID:g1195555
R;Shi, G.P.; Chapman, H.A.; Bhairi, S.M.; DeLeeuw, C.; Reddy, V.Y.; Weiss, S.J.
FEBS Lett. 357, 129-134, 1995
A;Title: Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of
A;Reference number: I38752; MUID:95104457; PMID:7805878
A;Accession: S68459
A;Molecule type: mRNA
A;Residues: 1-329 <SHI>
A;Cross-references: EMBL:U13665; NID:g606922; PIDN:AAA65233.1; PID:g606923
A;Experimental source: tissue-type blood
C;Genetics:
A;Gene: GDB:CTSK; PYCD; CTSO1; CTSO2; PKND; CTSO
A;Cross-references: GDB:453910; OMIM:265800; OMIM:601105
A;Map position: Iq21-Iq21
C;Superfamily: papain
C;Keywords: cysteine proteinase; glycoprotein; hydrolase
F;1-15/Domain: signal sequence #status predicted <SIG>
F;116-114/Domain: propeptide #status predicted <PRO>
F;115-329/Product: cathepsin K #status predicted <MAT>
F;103,161,213/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;139,276,296/Active site: Cys, His, Asn #status predicted

Query Match      26.8%; Score 323.5; DB 2; Length 329;
Best Local Similarity 36.0%; Pred. No. 1.6e-21;
Matches 73; Conservative 30; Mismatches 95; Indels 5; Gaps 5;

QY 8 GNAPAEIDLQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAQELVDCAS 67
DB 113 GRAPDSVDYRKGYVTPVKNQGGCGSCWAFSSVGALEGQLKTKTLLNLSQNLVDCVS 172
QY 68 QH-CHGDTIPRGTEYIQHN-GVQESYRYVAREQSCRRPNAQRFGISNYCOIYPPNA 124
DB 173 ENDCCGGYMNAYEYLKHNGLTESYTPYQAVGPGCGRLAIYAKVGYTVVHSGDEIE 232
QY 125 NKIREALQTHSAIAVIGIKDLDAFRHYDGRITIORDNQYQPNYHVNITVGSNAQVD 184
DB 233 KALKRAVARV-GPVSAVID-ASLTSFQYSGVYVDESCNNDNLNHAFLAVLAGYGIQGNK 290
QY 185 YWIVRNSWDTNWGDNGYGYFAAN 207
DB 291 HWIIRNSWGNWGNKGYILMARN 313

RESULT 11
T03941
cysteine proteinase (EC 3.4.22.-) precursor - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T03941
R;Becker, C.; Muentz, K.
submitted to the EMBL Data Library, September 1997
A;Description: cDNA cloning of a CPRI-homologous proteinase from germinating tobacco seed
```

A:Reference number: Z15148  
A:Accession: T03941  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-374 <BEC>  
A:CROSS-references: UNIPROT:Q24137; EMBL:299173; PIDN:CAB16317.1  
A:Experimental source: clone TCRP1  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-374/Product: cysteine proteinase #status predicted <MAT>  
F:164,299,319/Active site: Cys, His, Asn #status predicted

Query Match 26.7%; Score 322.5; DB 2; Length 374;  
Best Local Similarity 36.8%; Pred. No. 2.3e-21;  
Matches 81; Conservative 28; Mismatches 92; Indels 19; Gaps 9;

QY 11 FAEDILRWRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDC--ASQ 68  
DB 141 PHSVDWRKRGVAPIKNGSCGSCWAFSTVAVEGINQIVTGEMITLSEQLVDCDRVON 200  
QY 69 HCHGDTTPRGTEYTHNGVVO-ESYRYVAREQSCR--RNAQRFGISNCOIYPPNAN 125  
DB 201 SCNGGLMDYAFETIISNGGMDTEKHPYRGVEGCDPVRKNYKVSDGGYEDV-PRNER 259  
QY 126 KIREALQTHSAIAVILIGIKDLDAFRHYDGRITIIQRDNGYQPNHAVNIGVSNAGQVDY 185  
DB 260 ALQKAVA--HQFVCVALEASG-RAFQLYSS-GVFTGEGEEVD-HGVVVVGSGEDGDY 314  
QY 186 WIVRNSWDTNWDNGYGYFAANI-----DLMMIEYYP 217  
DB 315 WIVRNSWGTGKNGYGYKVRNRYKXSHLKGKOGIMTEASYP 354

RESULT 12  
C86413  
cysteine proteinase (EC 3.4.22.-) [similarity] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: C86413  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huitzer, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C86413  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-346 <STO>  
A:CROSS-references: UNIPROT:Q9LP39; GB:AE005172; NID:G9502427; PIDN:AAF98126.1; GSPDB:GN  
C:Genetics:  
A:Map position: 1  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:154,289,310/Active site: Cys, His, Asn #status predicted

Query Match 26.7%; Score 322; DB 2; Length 346;  
Best Local Similarity 32.3%; Pred. No. 2.4e-21;  
Matches 74; Conservative 33; Mismatches 86; Indels 36; Gaps 7;

QY 1 TNACISNGNAPAEI-----DLRQWRTVTPIRMGGCGSCWAFSGV 40  
DB 101 TGLRGVNTSPPEVNVNTPKPNWTVSDVLGTNKKWRNEGATPVKSGEGCGGWAFAI 160  
QY 41 AATESAYLAHRNQSLLDAEQELVDCASQ--HCHGDTTPRGIEY-IQNGVVOESYRYV 97  
DB 161 AAVEGLTKIARGNLISLSQQLDTRQNCKGTFVNAFNFIKRGISSENEYPYQ 220

QY 98 AREQSCRRPNAQRFGISNCOIYPPNANKIREMALQTHSAIAVILIGIKOLDA----FRHY 153  
DB 221 VKEGPCRSNARPAAILRGFENFVSNRERALLEAVSQPVAVAI-----DASEAGFVHY 273  
QY 154 DGRITIIQRDNGYQPNHAVNIGVY-SNAQVDYVIVRNSWDTNWDNGNGY 201  
DB 274 SGGVYNARNGCTSVN-HAVTLVGYGTSPEGMKYLAKNSGKWTGENGY 321

RESULT 13  
S12581  
cysteine proteinase (EC 3.4.22.-) precursor - black gram  
N:Alternate names: cysteine endopeptidase; sulphydryl endopeptidase  
C:Species: Vigna mungo (black gram)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S12581; S05497; S20213; S48684  
R:Akasofu, H.; Yamauchi, D.; Minamikawa, T.  
Nucleic Acids Res. 18, 1892, 1990  
A:Title: Nucleotide sequence of the gene for the Vigna mungo sulphydryl-endopeptidase (st  
A:Reference number: S12581; MUID:90245586; PMID:2336365  
A:Accession: S12581  
A:Molecule type: DNA  
A:Residues: 1-362 <AKAL>  
A:CROSS-references: UNIPROT:P12412; EMBL:X51900; NID:G22065; PIDN:CAA36181.1; PID:G22066  
R:Akasofu, H.; Yamauchi, D.; Mitsuhashi, W.; Minamikawa, T.  
Nucleic Acids Res. 17, 6733, 1989  
A:Title: Nucleotide sequence of cDNA for sulphydryl-endopeptidase (SH-EP) from cotyledons  
A:Reference number: S05497; MUID:89386007; PMID:2780300  
A:Accession: S05497  
A:Molecule type: mRNA  
A:Residues: 1-362 <AKA2>  
A:CROSS-references: EMBL:X15732; NID:G22061; PIDN:CAA33753.1; PID:G22062  
A:Accession: S20213  
A:Molecule type: protein  
R:Residues: 132-140 <AKA3>  
R:Okamoto, T.; Nakayama, H.; Seta, K.; Ise, T.; Minamikawa, T.  
FEBS Lett. 351, 31-34, 1994  
A:Title: Posttranslational processing of a carboxy-terminal propeptide containing a KDEI  
A:Reference number: S48684; MUID:94357273; PMID:8076688  
A:Accession: S48684  
A:Molecule type: protein  
A:Residues: 127-140/197-216;324-333;339-352 <OKA>  
C:Genetics:  
A:Introns: 150/1; 228/3; 275/3  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-362/Product: cysteine proteinase #status predicted <MAT>  
F:152,288,309/Active site: Cys, His, Asn #status predicted

Query Match 26.7%; Score 322; DB 1; Length 362;  
Best Local Similarity 35.0%; Pred. No. 2.5e-21;  
Matches 78; Conservative 33; Mismatches 94; Indels 18; Gaps 8;

QY 8 GNAPAEIDLQMETVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQELVDCAS 67  
DB 126 GSVFASVDWRKRGAVTDVKDQGGCGSCWAFSTIVAVEGINQIKTNKVLSEQLVDCDK 185  
QY 68 Q--HGCHGDTTPRGIEYI-OHNGVVOESYRYVAREQSC--RRPNAQRFGISNCOIYPP 122  
DB 186 EENQCGNGGLMESAFEFIKQGGITTESNYPYTAQSGTCDESKVNDLAVSIDGHENVPVN 245  
QY 123 NANKIREALQTHSAIAVILIGIKDLDAFRHYDGRITIIQRDNGYQPNHAVNIGVY-SNAQ 181  
DB 246 DENALLKAVANQPVSVDAIDAGSD---FQFY-SEGFTGDCNTDLN-HGVAIVGYGTVD 300  
QY 182 GVDYVIVRNSWDTNWDNGYGYFAANID-----LMMIEYYP 217  
DB 301 GTNVIIVRNSWGTGKNGYGYKVRNRYKXSHLKGKOGIMTEASYP 343

RESULT 14



OM protein - protein search, using sw model

Run on: October 27, 2004, 17:34:27 ; Search time 328.135 Seconds  
(without alignments)  
389.270 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSIGNAPAEIDLQRM.....YFAANIDLMTEEPYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1196	99.2	320	1	MMAL DERP	P08176 dermatophag
2	1037.5	86.0	321	1	EUM1_EURMA	P25780 euroglyphus
3	1005.5	83.4	321	2	BAC53948	BAC53948 dermatoph
4	997.5	82.7	321	1	MMAL DERP	P16311 dermatophag
5	944.5	78.3	210	2	Q9GY0	Q9GY0 dermatophag
6	607	50.3	146	2	Q95X05	Q95X05 dermatophag
7	591.5	49.0	263	2	Q819P1	Q819P1 psoroptes o
8	469	38.9	107	2	Q95X04	Q95X04 dermatophag
9	437	36.2	133	2	Q958Y3	Q958Y3 dermatophag
10	389.5	32.3	94	2	Q7M431	Q7M431 dermatophag
11	382	31.7	221	2	Q95P04	Q95P04 blomia trop
12	341.5	28.3	333	2	Q75XQ7	Q75XQ7 petromyzon
13	332.5	27.6	366	2	Q6ZHP9	Q6ZHP9 oryza sativ
14	332.5	27.6	366	2	BAD09165	BAD09165 oryza sat
15	332	27.5	348	2	Q9ZQH7	Q9ZQH7 arabidopsis
16	331	27.4	326	2	Q717S6	Q717S6 callosobr
17	331	27.4	326	2	AQ11969	AQ11969 callosobr
18	330.5	27.4	393	2	Q6E7B4	Q6E7B4 brugia mala
19	328.5	27.2	326	2	Q67857	Q67857 fasciola gi
20	328.5	27.2	326	2	AAR08900	AAR08900 fasciola
21	328.5	27.2	460	2	Q7XXU7	Q7XXU7 anthurium a
22	328	27.2	362	2	Q7X7A6	Q7X7A6 glycine max
23	327.5	27.2	395	1	CATL_BRUPA	Q17473 brugia pata
24	327	27.1	348	1	PAP3_CARPA	P10056 carica papa
25	327	27.1	367	2	Q42673	Q42673 carica papa
26	326.5	27.1	326	2	Q42494	Q42494 fasciola he
27	326.5	27.1	327	2	Q6QXF0	Q6QXF0 agrotis seg
28	326.5	27.1	327	2	AAS82707	AAS82707 agrotis s
29	326	27.0	326	2	Q717S5	Q717S5 callosobr
30	326	27.0	326	2	AQ11970	AQ11970 callosobr
31	325	26.9	326	2	Q717S9	Q717S9 callosobr

AAQ11966 callosobr  
Q717S8 fasciola he  
Q717C0 callosobr  
AAQ11965 callosobr  
Q9Xf79 sandersonia  
P43235 homo sapien  
P61276 macaca fasc  
P61277 macaca mulla  
Q6fhn2 homo sapien  
Q86gf7 pandalus bo  
Q717S1 callosobr  
AAQ11974 callosobr  
Q93XQ9 ipomoea bat  
Q717S4 callosobr

#### ALIGNMENTS

RESULT 1

ID	MMAL DERP	STANDARD;	PRT;	320 AA.
AC	P08176; Q24616;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Major mite fecal allergen Der p 1 precursor (EC 3.4.22.-) (Der p I).			
GN	Name=DERP1;			
OS	Dermatophagoides pteronyssinus (House-dust mite).			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
OC	Acariformes; Setaformis; Astigmata; Psoroptidia; Analgoidea;			
OC	Pyroglyphidae; Dermatophagoides.			
OX	NCBI_taxID=8956;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND POLYMORPHISM.			
RP	MEDLINE=9357682; PubMed=8353459;			
RA	Chua K.Y., Kehal P.K., Thomas W.R.;			
RT	"Sequence polymorphisms of cDNA clones encoding the mite allergen Der			
RT	p.i."			
RL	Int. Arch. Allergy Immunol. 101:364-368(1993).			
RN	[2]			
RP	SEQUENCE OF 76-320 FROM N.A.			
RP	MEDLINE=88089411; PubMed=3335830;			
RA	Chua K.Y., Stewart G.A., Thomas W.R., Simpson R.J., Dilworth R.J.,			
RA	Plozza T.M., Turner K.J.;			
RT	"Sequence analysis of cDNA coding for a major house dust mite			
RT	allergen, Der p.i. Homology with cysteine proteases."			
RL	J. Exp. Med. 167:175-182(1988).			
RN	[3]			
RP	SEQUENCE OF 81-176 FROM N.A.			
RP	MEDLINE=88114080; PubMed=3276629;			
RA	Thomas W.R., Stewart G.A., Simpson R.J., Chua K.Y., Plozza T.M.,			
RA	Dilworth R.J., Nisbet A., Turner K.J.;			
RT	"Cloning and expression of DNA coding for the major house dust mite			
RT	allergen Der p.i. in Escherichia coli."			
RL	Int. Arch. Allergy Appl. Immunol. 85:127-129(1988).			
RN	[4]			
RP	REVIEWS TO 232-241.			
RP	MEDLINE=91215493; PubMed=2021874;			
RA	Dilworth R.J., Chua K.Y., Thomas W.R.;			
RT	"Sequence analysis of cDNA coding for a major house dust mite			
RT	allergen, Der p.i."			
RL	Clin. Exp. Allergy 21:25-32(1991).			
RN	[5]			
RP	SEQUENCE OF 99-308 FROM N.A.			
RP	MEDLINE=93130112; PubMed=1483062;			
RA	Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;			
RT	"Molecular characterisation of Group I allergen Eur m I from house			
RT	dust mite Euroglyphus maynei."			
RL	Int. Arch. Allergy Immunol. 99:150-152(1992).			
RN	[6]			
RP	SEQUENCE OF 99-127.			
RP	MEDLINE=88229138; PubMed=3372999;			

RA Lind P., Hansen O.C., Horn N.;  
 RT "the binding of mouse hybridoma and human IgE antibodies to the major  
 RT fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative  
 RT binding site location and species specificity studied by solid-phase  
 RT inhibition assays with radiolabeled antigen.";  
 RL J. Immunol. 140:4256-4262(1988).  
 RN [7]  
 RP SEQUENCE OF 99-139; 177-192; 208-224 AND 260-277, AND VARIANT ALA-222.  
 RX PubMed-2911558;  
 RA Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;  
 RT "Structural studies on the allergen Der p1 from the house dust mite  
 RT Dermatophagoides pteronyssinus: similarity with cysteine  
 RT proteinases.";  
 RL Protein Seq. Data Anal. 2:17-21(1989).  
 RN [8]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE=95062135; PubMed=7971950;  
 RA Topham C.M., Srinivasan N., Thorpe C.J., Overington J.P.,  
 RA Kalsheker N.A.;  
 RT "Comparative modelling of major house dust mite allergen Der p 1:  
 RT structure validation using an extended environmental amino acid  
 RT propensity table.";  
 RL Protein Eng 7:869-894(1994).  
 CC -!- FUNCTION: Thiol protease that hydrolyzes proteins, with a  
 CC preference for Phe or basic residues.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
 CC mite allergy are bronchial asthma, allergic rhinitis and  
 CC conjunctivitis. Reacts with IgE in 80% of patients with house dust  
 CC allergy.  
 CC -!- SIMILARITY: Belongs to peptidase family C1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U11695; AAB60215.1; -;  
 DR EMBL; M24794; AAA28296.1; ALT\_INIT.  
 DR EMBL; X65197; CAA45317.1; -;  
 DR PIR; JQ0337; JQ0337.  
 DR HSP; P53634; IK3B.  
 DR MEROPS; C01.073; -;  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; Pept\_cys\_actsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPA1N.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SM00845; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIO1\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIO1\_PROTEASE CYS; 1.  
 DR PROSITE; PS00639; THIO1\_PROTEASE HIS; 1.  
 KW Allergen, Direct protein sequencing; Glycoprotein; Hydrolase;  
 KW Polymorphism; Signal; Thiol protease; Zymogen.  
 FT SIGNAL 1 18  
 FT PROPEP 19 98  
 FT CHAIN 99 320  
 FT CARBOHYD 150 150  
 FT ACT\_SITE 132 132  
 FT ACT\_SITE 268 268  
 FT ACT\_SITE 288 288  
 FT ACT\_SITE 288 288  
 FT DISULFID 102 215  
 FT DISULFID 129 169  
 FT DISULFID 163 201  
 FT VARIANT 148 148  
 FT VARIANT 179 179  
 FT VARIANT 222 222  
 FT VARIANT 234 234  
 FT VARIANT 313 313  
 FT SEQUENCE 320 AA; 36104 MW; A0B1F4DD09791DFE CRC64;  
 SO

Query Match 99.2%; Score 1196; DB 1; Length 320;  
 Best Local Similarity 99.1%; Pred. No. 3.9e-99;  
 Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLQMRVTPTPIRMQGGCGGSCWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 DB 99 TNACISNGNAPAEIDLQMRVTPTPIRMQGGCGGSCWAFSGVAATESAYLAHRNOSLDLAEQ 158  
 QY 61 ELVDCASQHGCHGTIPRGIEYIOHNGVQVESYRYVAREQSCRRPNAQRFGLSNVCQIY 120  
 DB 159 ELVDCASQHGCHGTIPRGIEYIOHNGVQVESYRYVAREQSCRRPNAQRFGLSNVCQIY 218  
 QY 121 PPNANKIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHNAVIVGYSNA 180  
 DB 219 PPNVKNKIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHNAVIVGYSNA 278  
 QY 181 QGVDTWIVRNSWDTNWDGNGYGFPAANIDLMIMEEYFYVWIL 222  
 DB 279 QGVDTWIVRNSWDTNWDGNGYGFPAANIDLMIMEEYFYVWIL 320  
 RESULT 2  
 EMBL\_EURMA STANDARD; PRT; 321 AA.  
 ID EMBL\_EURMA Q9TZZ4; Q9TZZ4; Q9UBA0;  
 AC F25780; Q9TZZ4; Q9TZZ4; Q9UBA0;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Mite group 1 allergen Eur m 1 precursor (EC 3.4.22.-) (Eur m 1).  
 GN Name=EURM1;  
 OS Eurygllyphus maynei (Mayne's house dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Eurygllyphus.  
 OX NCBI\_taxID=6958;  
 RN [1]  
 RP SEQUENCE FROM N.A. (EUR M 1.0101 AND EUR M 1.0102).  
 RX MEDLINE=99126275; PubMed=9925958;  
 RA Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;  
 RT "Molecular analysis of the group 1 and 2 allergens from the house dust  
 RT mite, Eurygllyphus maynei.";  
 RL Int. Arch. Allergy Immunol. 118:15-22(1999).  
 RN [2]  
 RP SEQUENCE OF 99-309 FROM N.A.  
 RX MEDLINE=93130112; PubMed=1483062;  
 RA Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
 RT "Molecular characterisation of group 1 allergen Eur m 1 from house  
 RT dust mite Eurygllyphus maynei.";  
 RL Int. Arch. Allergy Immunol. 99:150-152(1992).  
 CC -!- FUNCTION: Probable thiol protease.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
 CC mite allergy are bronchial asthma, allergic rhinitis and  
 CC conjunctivitis.  
 CC -!- SIMILARITY: Belongs to peptidase family C1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF047610; AAC82351.1; -;  
 DR EMBL; AF047611; AAC82352.1; ALT\_INIT.  
 DR EMBL; AF047612; AAC82353.1; -;  
 DR EMBL; X60073; CAA42677.1; -;  
 DR PIR; S21864; S21864.  
 DR HSP; P53634; IK3B.  
 DR MEROPS; C01.073; -;  
 DR InterPro; IPR000668; Peptidase\_C1.  
 SO



DR InterPro; IPR000169; Pept\_cys\_acsite.  
DR Pfam; PF00112; Peptidase\_C1; I.  
DR PRINTS; PR00705; PAPAIN\_C1.  
DR ProDom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00640; THIOLESTERASE ASN; 1.  
DR PROSITE; PS00139; THIOLESTERASE CYS; 1.  
DR PROSITE; PS00639; THIOLESTERASE HIS; 1.  
KW Allergen; Glycoprotein; Hydrolase; Signal; Thiol protease; Zymogen.  
FT SIGNAL 1 18 Potential.  
FT PROPEP 19 98  
FT CHAIN 99 321 Mite group 1 allergen Eur m 1.  
FT ACT\_SITE 133 133 By similarity.  
FT ACT\_SITE 269 269 By similarity.  
FT ACT\_SITE 289 289 By similarity.  
FT DISULFID 130 170 By similarity.  
FT CARBOHYD 34 34 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).  
FT VARIANT 36 36 T -> S (in Eur m 1.0102).  
FT VARIANT 126 126 M -> N (in Eur m 1.0102).  
FT VARIANT 320 320 M -> I (in Eur m 1.0102).  
SQ SEQUENCE 321 AA; 36290 MW; 6CFD44FEC725999E CRC64;  
  
Query Match 86.0%; Score 1037.5; DB 1; Length 321;  
Best Local Similarity 83.9%; Pred. No. 6.9e-85;  
Matches 187; Conservative 19; Mismatches 16; Indels 1; Gaps 1;  
  
QY 1 TNACSG-NAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAE 59  
Db 99 TVACINSVSLPSELDLSRLVTPIRMGGCGSCWAFSGVASTESAYLAHRNOSLDLAE 158  
QY 60 QELVDCASQHGCHGTIPRGIEYIQHNGVQVSEYRYVARQSCRRPNAQRFGISNYCQI 119  
Db 159 QELVDCASQHGCHGTIPRGIEYIQHNGVQVSEYRYVARQSCRRPNAQRFGISNYCQI 218  
QY 120 YPNANKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSN 179  
Db 219 SPDSNKIRQALTQHTAVAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSN 278  
QY 180 AGVDYWIVRNSWDITWNGDNGYGYFAANIDLMWIEEYPPYVIL 222  
Db 279 TQGVYWIVRNSWDITWNGDNGYGYFAANIDLMWIEEYPPYVIM 321  
  
RESULT 3  
BACS3948 PRELIMINARY; PRT; 321 AA.  
AC BACS3948;  
DT 02-WAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-WAR-2004 (TrEMBLrel. 27, Last sequence update)  
DE Der f 1 allergen preproenzyme precursor.  
OS Dermatophagoides farinae (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;  
OC Dermatophagoides.  
OX NCBI\_TaxID=6954;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yasuhara T., Takai T., Takahashi K., Yuuki T., Yokota T., Okudaira H.,  
RA Okumura Y.;  
RT "Dermatophagoides farinae Der f 1 allergen preproenzyme mRNA.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Dilworth R.J., Chuan K.Y., Thomas W.R.;  
RT "Sequence analysis of cDNA coding for a major house dust mite  
allergen, Der f 1.";  
RL Clin. Exp. Allergy 21:25-32 (1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93357682; PubMed=8353459;  
RA Chua K.Y., Kehai P.K., Thomas W.R.;

RT "Sequence polymorphisms of cDNA clones encoding the mite allergen Der  
p 1.";  
RL Int. Arch. Allergy Immunol. 101:364-368 (1993).  
DR EMBL; AB034946; BACS3948.1; -.  
KW Signal.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 99 321  
FT VARIANT 201 201 Q -> R (IN REF. 2).  
FT VARIANT 282 282 V -> D (IN REF. 2).  
SQ SEQUENCE 321 AA; 36391 MW; 83594754EEBB4477 CRC64;  
  
Query Match 83.4%; Score 1005.5; DB 2; Length 321;  
Best Local Similarity 81.6%; Pred. No. 5.2e-82;  
Matches 182; Conservative 18; Mismatches 22; Indels 1; Gaps 1;  
  
QY 1 TNACSG-NAPAEIDLROMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAE 59  
Db 99 TSACRINSVNPSELDRSLRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNOSLDLAE 158  
QY 60 QELVDCASQHGCHGTIPRGIEYIQHNGVQVSEYRYVARQSCRRPNAQRFGISNYCQI 119  
Db 159 QELVDCASQHGCHGTIPRGIEYIQHNGVQVSEYRYVARQSCRRPNAQRFGISNYCQI 218  
QY 120 YPNANKIREALQTHSAIAVIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSN 179  
Db 219 YPDPVKQIREALTQHTAVAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSN 278  
QY 180 AGVDYWIVRNSWDITWNGDNGYGYFAANIDLMWIEEYPPYVIL 222  
Db 279 TQGVYWIVRNSWDITWNGDNGYGYFAANIDLMWIEEYPPYVIM 321  
  
RESULT 4  
MMAL DERFA STANDARD; PRT; 321 AA.  
AC PL631L;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Major mite fecal allergen Der f 1 precursor (EC 3.4.22.-) (Der f I).  
GN Name=DERf1;  
OS Dermatophagoides farinae (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptida; Analgoidea;  
OC Pyroglyphidae; Dermatophagoides.  
OX NCBI\_TaxID=6954;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dilworth R.J., Chua K.Y., Thomas W.R.;  
RT "Sequence analysis of cDNA coding for a major house dust mite  
allergen, Der f 1.";  
RL Clin. Exp. Allergy 21:25-32 (1991).  
RN [2]  
RP SEQUENCE OF 98-309 FROM N.A.  
RA Kent N., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 99-128.  
RX MEDLINE=88291938; PubMed=3372999;  
RA Lind P., Hansen O.C., Horn N.;  
RT "The binding of mouse hybridoma and human IgE antibodies to the major  
fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative  
binding site location and species specificity studied by solid-phase  
inhibition assays with radiolabeled antigen.";  
RL J. Immunol. 140:4256-4262 (1988).  
CC -!- FUNCTION: Thiol protease that hydrolyzes proteins, with a  
preference for phe or basic residues.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
mite allergy are bronchial asthma, allergic rhinitis and  
conjunctivitis.  
CC -!- SIMILARITY: Belongs to peptidase family C1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X65196; CAA46316.1; -  
 DR PIR; A27634; A27634.  
 DR HSP; P53634; IK3B.  
 DR MROFS; C01.073; -  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; Pept\_cys\_actsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPA1N.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SM00645; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; 1.  
 KW Allergen. Direct protein sequencing; Glycoprotein; Hydrolase; Signal;  
 KW Thiol protease; Zymogen.  
 FT SIGNAL 1 18 Potential.  
 FT PROPEP 19 98 Activation peptide.  
 FT CHAIN 99 321 Major mite fecal allergen Der f 1.  
 FT ACT\_SITE 133 133 By similarity.  
 FT ACT\_SITE 269 269 By similarity.  
 FT ACT\_SITE 288 288 By similarity.  
 FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).  
 FT DISULFID 102 216 By similarity.  
 FT DISULFID 130 170 By similarity.  
 FT DISULFID 164 202 By similarity.  
 FT CONFLICT 201 201 R -> Q (in Ref. 2).  
 FT CONFLICT 282 282 D -> V (in Ref. 2).  
 SQ SEQUENCE 321 AA; 36435 MW; 04523E54EBB476E CRC64;

Query Match 82.7%; Score 997.5; DB 1; Length 321;  
 Best Local Similarity 81.2%; Pred. No. 2.7e-81;  
 Matches 181; Conservative 16; Mismatches 23; Indels 1; Gaps 1;

QY 1 TNACISNG-NAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHNSQLDLAE 59  
 Db 99 TSACRINSVNPSELDLSRLTTPPIRMQGGCGSWAFSGVAATESAYLAHNSQLDLSE 158  
 QY 60 QELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCYI 119  
 Db 159 QELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCYI 218  
 QY 120 YPPNANKIRALQTHSAIAVIIGIKDLAPFHYDGTIIQHDNGYOPNYHVNIVGYSN 179  
 Db 219 YPPDVVKIRALQTHSAIAVIIGIKDLAPFHYDGTIIQHDNGYOPNYHVNIVGYSN 278  
 QY 180 AQGVYDWIVRNSWDTTWGDSGYGFAGNIDLMIEEPPYVWIL 222  
 Db 279 TQGDYDWIVRNSWDTTWGDSGYGFAGNIDLMIEEPPYVWIM 321

RESULT 5  
 Q9GYO PRELIMINARY; PRT; 210 AA.  
 ID Q9GYO  
 AC Q9GYO  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Allergen Der f1 (Fragment).  
 OS Dermatophagoides farinae (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Dermatophagoides.  
 OC NCBI\_TaxID=6954;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Hao M.Q., Xu J., Zhong N.S.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to peptidase family C1.  
 DR EMBL; AF285763; AAG00520.1; -  
 DR PIR; A27634; A27634.  
 DR GO; GO:0004137; F:cysteine-type endopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; Pept\_cys\_actsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPA1N.  
 DR SMART; SM00645; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; UNKNOWN\_1.  
 KW Hydrolase; Protease; Thiol protease.  
 FT NON\_TER 1 210  
 FT NON\_TER 210 210  
 SQ SEQUENCE 210 AA; 23548 MW; BA08029D642EEB90 CRC64;

Query Match 78.3%; Score 944.5; DB 2; Length 210;  
 Best Local Similarity 81.4%; Pred. No. 9.6e-77;  
 Matches 171; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

QY 2 NACISNG-NAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHNSQLDLAEQ 60  
 Db 1 SACRINSVNPSELDLSRLTTPPIRMQGGCGSWAFSGVAATESAYLAHNSQLDLSE 60  
 QY 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCYI 120  
 Db 61 ELVDCASQHGCHGDTPIRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCYI 120  
 QY 121 PPNANKIRALQTHSAIAVIIGIKDLAPFHYDGTIIQHDNGYOPNYHVNIVGYSNA 180  
 Db 121 PPDVVKIRALQTHSAIAVIIGIKDLAPFHYDGTIIQHDNGYOPNYHVNIVGYSN 180  
 QY 181 QGVYDWIVRNSWDTTWGDSGYGFAGNIDL 210  
 Db 181 QGVYDWIVRNSWDTTWGDSGYGFAGNIDL 210

RESULT 6  
 Q9SX05 PRELIMINARY; PRT; 146 AA.  
 ID Q9SX05  
 AC Q9SX05  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cysteine proteinase (Fragment).  
 GN Name=CPW2.  
 OS Dermatophagoides farinae (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Dermatophagoides.  
 OC NCBI\_TaxID=6954;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Park H., Park S.Y., Kim K.Y., Park S.K., Yun H.C.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF194431; AAL14424.1; -  
 DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 FT NON\_TER 1 146  
 FT NON\_TER 146 146  
 SQ SEQUENCE 146 AA; 16852 MW; BB304800946D4047 CRC64;

Query Match 50.3%; Score 607; DB 2; Length 146;  
 Best Local Similarity 84.2%; Pred. No. 1.4e-46;  
 Matches 112; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 39 GVAATESAYLAHRNOSLDLAEQELVDCASQHGCHGDTIPRGIEYTOHNGVVOESYRVA 98  
Db 1 GVAATESAYLAHRNOSLDLAEQELVDCASQHGCHGDTIPRGIEYTOHNGVVOESYRVA 60  
QY 99 REOSCRPNNAORFGISNYCQIYPPNANKIREALAOHSAIAVIIGIKDLDAFRHYDGRGTI 158  
Db 61 REOQCRPNNSOHYGSNYCQIYPPVVKQIREALTQTHAIAVIIGIKDLRAFOHYDGRGTI 120  
QY 159 IORDNGYQNYHA 171  
Db 121 IOHNGYQNYHA 133  
RESULT 7  
Q819P1 PRELIMINARY; PRT; 263 AA.  
ID Q819P1  
AC Q819P1  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Derp1 antigen (Fragment).  
OS Psoroptes ovis (Sheep scab mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Sarcotoidae;  
OC Psoroptidae; Psoroptes.  
OX NCBI\_TaxID=83912;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22294898; PubMed=12406195;  
RA Lee A.J., Machel J., Van Den Broek A.H.M., Nisbet A.J.,  
RA Miller H.R.P., Isaac R.E., Huntley J.F.;  
RT "Identification of an antigen from the sheep scab mite, Psoroptes  
RT ovis, homologous with house dust mite group I allergens";  
RL Parasite Immunol. 24:413-422(2002).  
DR EMBL; AF495854; AA014671.1; -.  
DR MEROPS; C01.073; -.  
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase C1.  
DR InterPro; IPR000169; Pept cys acsite.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR Prodom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
FT NON\_TER 1  
FT NON\_TER 263  
SQ SEQUENCE 263 AA; 29576 MW; BF6DD21006DAB5B0 CRC64;  
Query Match 49.0%; Score 591.5; DB 2; Length 263;  
Best Local Similarity 64.4%; Pred. No. 6.8e-45;  
Matches 112; Conservative 22; Mismatches 39; Indels 1; Gaps 1;  
QY 3 ACSING-NAPADIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNOSLDLAEQ 61  
Db 90 ACQIGAVNIPNIDLRALGVTKIKNOVACGSCWAFSGVATVESNLYSDNVLDSQ 149  
QY 62 LVDCASQHGCHGDTIPRGIEYTOHNGVVOESYRVAEQSCRRPNAORFGISNYCQIYP 121  
Db 150 LVDCASQHGCGDTVLNGLRYIQKNGVVEQSYPYKAREGRQRENAKRYGKDLQIYP 209  
QY 122 PNANKIREALAOHSAIAVIIGIKDLDAFRHYDGRGTIORDNGYQPNHAVNIV 175  
Db 210 PNGDKIRTYLATKQALSVIIGIRDLSFRHYDGRGTIORDNGYQPNHAVNIV 263  
RESULT 8  
Q95X04 PRELIMINARY; PRT; 107 AA.  
ID Q95X04  
AC Q95X04  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cysteine proteinase (Fragment).

GN Name=CPW3;  
OS Dermatophagoides farinae (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;  
OC Pyroglyphidae; Dermatophagoides.  
OX NCBI\_TaxID=5954;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Park H., Park S.Y., Kim K.Y., Park S.K., Yun H.C.;  
RL Submitted (OCT-1999); to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF194432; AAL14425.1; -.  
DR InterPro; IPR000169; Pept cys acsite.  
DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 107  
SQ SEQUENCE 107 AA; 12277 MW; A80E7876CBA6F97A CRC64;  
Query Match 38.9%; Score 469; DB 2; Length 107;  
Best Local Similarity 80.4%; Pred. No. 2.4e-34;  
Matches 86; Conservative 8; Mismatches 13; Indels 0; Gaps 0;  
QY 77 PRGIEYTOHNGVVOESYRVAEQSCRRPNAORFGISNYCQIYPPNANKIREALAOHSA 136  
Db 1 PRGIEYTOHNGVVOESYRVAEQSCRRPNAORFGISNYCQIYPPVVKQIREALTQTH 60  
QY 137 AIAVIIGIKDLDAFRHYDGRGTIORDNGYQPNHAVNIVGYSNAQV 183  
Db 61 AIAVIIGIKDLRAFOHYDGRGTIIOHNGYQPNHAVNIVGISTQV 107  
RESULT 9  
Q968Y3 PRELIMINARY; PRT; 133 AA.  
ID Q968Y3  
AC Q968Y3  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Cysteine proteinase (Fragment).  
OS Dermatophagoides pteronyssinus (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;  
OC Pyroglyphidae; Dermatophagoides.  
OX NCBI\_TaxID=6956;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Park H., Yun H.C., Kim K.Y., Park S.Y., Park S.K.;  
RL Submitted (APR-1999); to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF145247; AAK38773.1; -.  
DR MEROPS; C01.073; -.  
DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase C1.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
FT NON\_TER 1  
FT NON\_TER 133  
SQ SEQUENCE 133 AA; 14965 MW; 5033C26B15E68B9C CRC64;  
Query Match 36.2%; Score 437; DB 2; Length 133;  
Best Local Similarity 61.4%; Pred. No. 2.3e-31;  
Matches 81; Conservative 17; Mismatches 34; Indels 0; Gaps 0;  
QY 31 CGSCWAFSGVAATESAYLAHRNOSLDLAEQELVDCASQHGCHGDTIPRGIEYTOHNGV 90  
Db 1 CGSCWAFAGVAATESAYLAHRNOSLDLAEQELVDCASQHGCHGDTIPRGIDYTOQNGIVE 60  
QY 91 ESYRYVAREQSCRRPNAORFGISNYCQIYPPNANKIREALAOHSAIAVIIGIKDLDA 150  
Db 61 EQAYEVNARENCEPPENPHSIEQYQDHSNVELIKTALDYKSAVAVINIHNNAF 120  
QY 151 RHYDGRGTIORD 162  
Db 121 RHYDGSYVITD 132

```

RESULT 10
Q7M431
ID Q7M431 PRELIMINARY; PRT; 94 AA.
AC Q7M431
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major fecal allergen Der p 1 (Fragments).
OS Dermatophagoides pteronyssinus (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6956;
RN [1]
RN R
RP MEDLINE=89098855; PubMed=2911558;
RX Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;
RA "Structural studies on the allergen Der p1 from the house dust mite
RT Dermatophagoides pteronyssinus: similarity with cysteine
RT proteinases.";
RL Protein Seq. Data Anal. 2:17-21 (1989).
DR PIR; S03380.
DR GO; 0008234; F:cysteine-type peptidase activity; IEA.
DR GO; 0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR ProDom; PD000158; Peptidase_C1; 1.
FT NON_TER 1 94
FT SEQUENCE 94 AA; 10327 MW; 98F744165C8428A8 CRC64;
SQ SEQUENCE 32.3%; Score 389.5; DB 2; Length 94;
Best Local Similarity 48.6%; Pred. No. 2.9e-27;
Matches 89; Conservative 0; Mismatches 4; Indels 89; Gaps 4;

QY 1 TNACSGNAPAEIDLRQMTVTPIR--MQGGCGSWAFSGVAATESAYLAHRNQSLLDLA 58
Db 1 TNACSGNAPAEIDLRQMTVTPIRQMGGCGSWAFSGVA----- 43
QY 59 EQLVDCASGCHGDDIPGIEYIQHGVVQSVYRVAREQSCRRPNAQFGISNYCQ 118
Db 44 -----GIEYIQHGVVQSVY-----RFGISNYCQ 68
QY 119 IYFPNANKIREALAQTHSAIAVIGIKDLDAFRHYDGRTHIQRDNGYQPNYHAVNIVGYS 178
Db 69 IYFPNANK-----DNGYQPNYXAVNIVGYX 93
QY 179 N 179
Db 94 N 94

RESULT 11
Q95PJ4
ID Q95PJ4 PRELIMINARY; PRT; 221 AA.
AC Q95PJ4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cysteine protease (Fragment).
OS Blomia tropicalis (Mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Glycyphagoidea;
OC Echinopodidae; Blomia.
OX NCBI_TaxID=40697;
RN [1]
RN R
RP SEQUENCE FROM N.A.
RX Mora C.I., Diaz A.M., Montealegre F., Flores I.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF277840; AAK58415.1; -.
DR GO; 0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; 0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.

```

```

DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR SMART; SMO0645; Pept_C1; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Protease.
FT NON_TER 1 221
FT CHAIN <1 221 cysteine protease.
SQ SEQUENCE 221 AA; 25126 MW; 272B45ESA53F2900 CRC64;
Query Match 31.7%; Score 382; DB 2; Length 221;
Best Local Similarity 38.4%; Pred. No. 3.8e-26;
Matches 86; Conservative 35; Mismatches 77; Indels 26; Gaps 8;

QY 11 PAEIDLRQMTVTPIRQMGGCGSWAFSGVAATESAYLAHRNQSLLDLAEOELVDCASQ-- 68
Db 2 PAFDWRQKTHVPIRQGGCGSWAFSGVAETLYAIHRHQNILSEQELLDCTHYLY 61
QY 69 -----EGCHGDTIPRGIEYIQHGVVQSVYRVAREQSCRRPNAQ--RFGISNYCQIY 120
Db 62 DPTYKCHGCGSGMSPFAFKYMKQKGLLESHYPYKMLNOC-QANARGTRYHYSSY---- 116
QY 121 PPNANKIREALAQTHSAI-----AVIGIKDLDA-FRHYDGRTHIQRDNGYQPNY--HAVN 173
Db 117 --NSLRYRAGDQETQAAIMNMHGPVVIYIHGTEAHFRL--RKGLRGAGYNDAQIDHAVY 172
QY 174 IVGVSNAGQVDYIVIRNSWDITNGDNGYGVFAANIDLMIEEYP 217
Db 173 LVGMGTQNGIDYIVIRTSWGTQMGDAGYGVFERHNSLGINNYP 216

RESULT 12
Q7SXQ7
ID Q7SXQ7 PRELIMINARY; PRT; 333 AA.
AC Q7SXQ7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cathepsin.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RN R
RP SEQUENCE FROM N.A.
RX MEDLINE=22871689; PubMed=14507309;
RA Uinuk-Ool T.S., Takezaki N., Kuroda N., Figueroa F., Sato A.,
RA Samonte I.E., Mayer W.E., Klein J.;
RA "Phylogeny of antigen-processing enzymes: cathepsins of a
RT cephalochordate, an agnathan and a bony fish.";
RL Scand. J. Immunol. 58:436-448 (2003).
RN [2]
RN R
RP SEQUENCE FROM N.A.
RX TISSUE=Typhlossole.
RA Uinuk-ool T.S., Takezaki N., Kuroda N., Figueroa F., Sato A.,
RA Samonte I.E., Mayer W.E., Klein J.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY333299; AAK01146.1; -.
DR GO; 0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; 0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; 1.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SMO0645; Pept_C1; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
SQ SEQUENCE 333 AA; 37500 MW; 08CD129A748B099A CRC64;
Query Match 28.3%; Score 341.5; DB 2; Length 333;
Best Local Similarity 36.0%; Pred. No. 2.7e-22;
Matches 80; Conservative 36; Mismatches 51; Indels 15; Gaps 9;

```

```

QY 9 NAPAEDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQELVDCASQ 68
DB 116 NLPQVQWRLKGYVTPVKEQGLCGSWAFSGVAATESAYLAHRNOSLDLAEQELVDCASQ 175
QY 69 H---CCHGDTIPRGIEY-IOHNGVVOESYRYRYVAREQSCR-RPNAQRFGISNYCOIYPPN 123
DB 176 YNNCGNGRSEALQYIIDNNGIDSELSYPEHADGKCRKPANVATKCSY-QFVEPS 234
QY 124 ANK--IREALAQTHSAIAVIIGIKDLAFRHYDGRITIIQRDNGYOPNYHVNIVGYSNAQ 181
DB 235 SNEEVLRQAVASV-GPIAIAAMN-ADLDTFKHYKSLFNEPSCDKSPN-HAMLVVGYGSL 291
QY 182 GVDYIVRNSWDTNMGDNGYGYFAANID-----LMMIEEYPVY 219
DB 292 GNDYIVRNSWDTNMGDNGYGYFAANID-----LMMIEEYPVY 333

RESULT 13
Q6ZHP9 PRELIMINARY; PRT; 366 AA.
AC Q6ZHP9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative cysteine proteinase.
GN Name=OJ1191.G08.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OJ1191.G08.11;"
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004047; BAD09165.1; -.
SQ SEQUENCE 366 AA; 40600 MW; 89529D2315608C4C CRC64;

Query Match 27.6%; Score 332.5; DB 2; Length 366;
Best Local Similarity 35.7%; Pred. No. 1.9e-21;
Matches 79; Conservative 34; Mismatches 91; Indels 17; Gaps 7;

QY 9 NAPAEDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQELVDCAS- 67
DB 142 NLPWVDWRKKGAVTPVKNQGGCGSWAFSTVAAVEGINQIVTGKLVSLSEQLMDCDNT 201
QY 68 -QHCHGDTIPRGIEYIOHN-GVVQESYRYRYVAREQSCR--RPNAQRFGISNYCOIYPPN 123
DB 202 FNHGCRGLMDFAFYIMGNQGIYTEDYPYLMEEGYCREKQPHSKVITITGYEDVPANS 261
QY 124 ANKIREALAQTHSAIAVIIGIKDLAFRHYDGRITIIQRDNGYOPNYHVNIVGYSNAQV 183
DB 262 ETSLLKALAHQPVSVGIAGSRD---FQYK-GIFDGCIGPD-HALTAVGYSYVG 316
QY 184 DYWIVRNSWDTNMGDNGY-----GYFAANIDLMMIEEYP 217
DB 317 DYIIMKNSGKNWGEQGYFRIRRGTKPGCVCDIYKIASYP 357

RESULT 15
Q9ZQH7 PRELIMINARY; PRT; 348 AA.
AC Q9ZQH7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cysteine proteinase.
GN Name=At2g27420;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.B., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrea A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

```

```

AC BAD09165;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative cysteine proteinase.
GN OJ1191.G08.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza; Oryza sativa.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OJ1191.G08.11;"
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004047; BAD09165.1; -.
SQ SEQUENCE 366 AA; 40600 MW; 89529D2315608C4C CRC64;

Query Match 27.6%; Score 332.5; DB 2; Length 366;
Best Local Similarity 35.7%; Pred. No. 1.9e-21;
Matches 79; Conservative 34; Mismatches 91; Indels 17; Gaps 7;

QY 9 NAPAEDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQELVDCAS- 67
DB 142 NLPWVDWRKKGAVTPVKNQGGCGSWAFSTVAAVEGINQIVTGKLVSLSEQLMDCDNT 201
QY 68 -QHCHGDTIPRGIEYIOHN-GVVQESYRYRYVAREQSCR--RPNAQRFGISNYCOIYPPN 123
DB 202 FNHGCRGLMDFAFYIMGNQGIYTEDYPYLMEEGYCREKQPHSKVITITGYEDVPANS 261
QY 124 ANKIREALAQTHSAIAVIIGIKDLAFRHYDGRITIIQRDNGYOPNYHVNIVGYSNAQV 183
DB 262 ETSLLKALAHQPVSVGIAGSRD---FQYK-GIFDGCIGPD-HALTAVGYSYVG 316
QY 184 DYWIVRNSWDTNMGDNGY-----GYFAANIDLMMIEEYP 217
DB 317 DYIIMKNSGKNWGEQGYFRIRRGTKPGCVCDIYKIASYP 357

RESULT 15
Q9ZQH7 PRELIMINARY; PRT; 348 AA.
AC Q9ZQH7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cysteine proteinase.
GN Name=At2g27420;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.B., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrea A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

```

RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]

RP SEQUENCE FROM N.A.

RA Yamada K., Bann J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,  
 RA Davis R.W., Ecker J.R., Theologis A.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to peptidase family C1.

DR EMBL; AC006232; AAL15594.1; -;  
 DR EMBL; AY064033; AAL36389.1; -;  
 DR EMBL; AY063388; AAM20029.1; -;  
 DR PIR; F84672; F84672;  
 DR HSP; P53634; I13B.  
 DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000668; Peptidase\_C1.

DR InterPro; IPR000169; Pept\_cys\_acsite.

DR Pfam; PF00112; Peptidase\_C1; 1.

DR PRINTS; PR00705; PAPAIN\_1.

DR ProDom; PD000158; Peptidase\_C1; 1.

DR SMART; SM00645; Pept\_C1; 1.

DR PROSITE; PS00640; THIOI PROTEASE ASN; 1.

DR PROSITE; PS00139; THIOI PROTEASE CYS; 1.

DR PROSITE; PS00639; THIOI PROTEASE HIS; UNKNOWN\_1.

KW Hydrolase; Protease; Thiol protease.

SQ SEQUENCE 348 AA; 38738 MW; E86ABEC2B553E76 CRC64;

Query Match

Best Local Similarity 27.5%; Score 332; DB 2; Length 348;

Matches 84; Conservative 32; Mismatches 86; Indels 30; Gaps 11;

QY 5 SINGNAEITDLQMRVTPIRMQGGCGWAFSGVAATESAYLAHNSLDLAEQLVD 64

Db 126 SDNGES--MDWRQEGAVTQVKTQGRGGCWAFSAVAEGITKTRGELVSLSEQQLD 182

QY 65 CASQH--GCHGDTIPRGIEY-IQHNGVQESYRYVAREQSCRRPNA-----QRFGISNY 116

Db 183 CDRDYNQCGRGGIMSKAFEYIKNQGITTEDNYPQESQQTCSSTLTSSPRAATISGY 242

QY 117 CQYPPNANKIREALQTHSAIAVIIGIKDLD-APRHYDGRITIIORDNGY--QPNYHAYN 173

Db 243 -ETVPMNN---EALLQAVSQQVSVGIEGTGAAFRHSYGGVF---NGECGTDLHAYT 294

QY 174 IVGYS-NAQGVYDVMVNSWDTNKGNGYGFANID-----LMVIEEYP 217

Db 295 IVGYSSECTKYVVKNSWGETWGENGYRIKRDVDAQCGGLAIFAYP 346

Search completed: October 27, 2004, 17:50:14

Job time : 331.135 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:26:32 ; Search time 306.024 Seconds  
(without alignments)  
260.234 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACISNGNAPAEIDLQMR.....YFAANIDLMIEEYVYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1206	100.0	222	5 AAO20568	Cysteine
2	1206	100.0	222	5 ABB98333	Cysteine
3	1206	100.0	245	2 AAR22433	Der p I a
4	1206	100.0	245	2 AAR39359	Dermatoph
5	1206	100.0	245	2 AAR51727	Der p I.
6	1206	100.0	245	2 AAR47063	Protein a
7	1206	100.0	245	2 AAR66545	Der p I a
8	1206	100.0	245	2 AAW71908	Dermatoph
9	1206	100.0	245	2 AAY50356	Dermatoph
10	1206	100.0	245	2 AAU18959	House dus
11	1206	100.0	245	5 ABG71812	European
12	1206	100.0	320	2 AAR49920	Protein a
13	1203	99.8	222	4 AAU07748	House dus
14	1200	99.5	222	4 AAB98347	D. pteron
15	1200	99.5	302	4 AAB98346	D. pteron
16	1200	99.5	302	5 AAM50823	Recombina
17	1200	99.5	320	8 ADK52158	Full leng
18	1199	99.4	362	8 ADM57302	Modular a
19	1196	99.2	222	4 AAU07746	House dus
20	1196	99.2	222	5 ABG67023	House dus
21	1196	99.2	222	8 ADK52140	Der p1 al
22	1196	99.2	223	4 AAB98345	D. pteron
23	1196	99.2	223	4 AAB98331	D. pteron
24	1196	99.2	282	8 ADL27398	Polypepti
25	1196	99.2	302	4 AAB98348	D. pteron

26	1196	99.2	302	4 AAB98330	D. pteron
27	1196	99.2	302	6 AAE36747	Dermatoph
28	1196	99.2	303	4 AAB98344	D. pteron
29	1196	99.2	320	2 AAY25580	D. pteron
30	1196	99.2	320	2 AAY08592	D. pteron
31	1196	99.2	320	4 AAB98329	D. pteron
32	1196	99.2	320	6 ABB80128	Der p1. 6
33	1196	99.2	320	6 ABB98482	Amino aci
34	1196	99.2	320	7 ADC34830	House dus
35	1196	99.2	320	7 ADE38098	European
36	1196	99.2	320	8 ADM57314	Modular a
37	1195	99.1	339	2 AAY08596	D. pteron
38	1191	98.8	320	6 ABU11147	House dus
39	1191	98.8	343	2 AAY08597	D. pteron
40	1189	98.6	222	4 AAU07747	House dus
41	1187	98.4	320	2 AAY08593	D. pteron
42	1186	98.3	320	2 AAY08595	D. pteron
43	1185	98.3	315	2 AAY08594	D. pteron
44	1184	98.2	302	6 AAE36748	Dermatoph
45	1184	98.2	302	6 AAE36749	Dermatoph

#### ALIGNMENTS

RESULT 1

AAO20568

ID AAO20568 standard; protein; 222 AA.

XX AC AAO20568;

XX DT 02-JAN-2003 (first entry)

XX DE Cysteine protease protein.

XX KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;

XX KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;

XX KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;

XX KW atopic eczema.

XX OS Dermatophagoides pteronyssinus.

XX PH Key Location/Qualifiers

FT Misc-difference 105 /note= "Encoded by ACC"

XX WO200278736-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-FR001098.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PR 03-MAY-2001; 2001FR-00005929.

XX PR 29-MAY-2001; 2001US-00867159.

XX PA (ANTI-) ANTIALIS SARL.

XX PI Loria E, Terrasse G, Trehin Y;

XX DR WPI; 2002-750636/81.

XX DR N-PSDB; AAL41281.

XX PT Antiallergic compositions containing an anti-histamine, a histamine

PT synthesis inhibitor, and optionally an allergen or nucleic acid coding

XX for the allergen.

XX PS Claim 13; Page 30-31; 32pp; French.

XX CC The invention relates to antiallergic compositions containing an anti-

CC histamine, a histamine synthesis inhibitor, and optionally an allergen or

CC isolated nucleic acid molecule that has at least one polynucleotide

CC sequence coding for the allergen, together with a pharmaceutical carrier.

CC The pharmaceutical composition of the invention is useful as a non-specific anti-allergic treatment, and also useful in the treatment of allergic hypersensitivity, allergic asthma, allergic rhinitis, and allergic and atopic eczema. This sequence represents the cysteine protease protein relating to the anti-allergic compositions of the invention

XX SQ Sequence 222 AA;

Query Match 100.0%; Score 1206; DB 5; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.5e-127;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSLNGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
DB 1 TNACSLNGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120

QY 121 PPNANKIREALAOQTHSAIAVIIGIKLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
DB 121 PPNANKIREALAOQTHSAIAVIIGIKLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180

QY 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMWTEEPYVIVL 222  
DB 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMWTEEPYVIVL 222

RESULT 2  
ABB98533  
ID ABB98533 standard; protein; 222 AA.

XX AC ABB98533;  
XX DT 13-DEC-2002 (first entry)  
XX DE Cysteine protease.  
XX KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
XX KW immunotherapy; allergen; allergic hypersensitivity reaction;  
XX KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
XX KW cysteine protease.

XX OS Dermatophagoides pteronyssinus.

XX PH Key Location/Qualifiers  
XX FT Misc-difference 105 /note= "Encoded by CCA"

XX FT FR2822709-A1.

XX PD 04-OCT-2002.

XX PF 03-MAY-2001; 2001FR-00005929.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PA (ANTI-) ANTIALIS SARL.

XX PI Loria E, Terrasse G, Trehin Y;

XX DR WPI; 2002-735037/80.

XX DR N-PSDB; ABQ80833.

XX PT Antiallergic composition, useful for preventing and treating e.g. asthma, rhinitis or eczema, containing at least two of allergen, antihistamine and histamine synthesis inhibitor.

XX PS Claim 7; Page 27-28; 33pp; French.

XX CC The present invention relates to an anti-allergic pharmaceutical

CC composition (I) comprising a pharmaceutical carrier containing an active agent combination of at least two of: an allergen; an antihistamine; and a histamine synthesis inhibitor. (I) is used for treating or preventing allergic hypersensitivity reactions, especially allergic asthma, allergic rhinitis or allergic atopic eczema, in babies, children or adults. The present sequence is cysteine protease from Dermatophagoides pteronyssinus, which was used as an allergen in the invention

XX SQ Sequence 222 AA;

Query Match 100.0%; Score 1206; DB 5; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.5e-127;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACSLNGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
DB 1 TNACSLNGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120

QY 121 PPNANKIREALAOQTHSAIAVIIGIKLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
DB 121 PPNANKIREALAOQTHSAIAVIIGIKLDAPRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180

QY 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMWTEEPYVIVL 222  
DB 181 QGVDTWIVRNSWDTNMGDNGYGYFAANIDLMWTEEPYVIVL 222

RESULT 3  
AAR22433  
ID AAR22433 standard; protein; 245 AA.

XX AC AAR22433;

XX DT 27-AUG-2003 (revised)

XX DT 04-AUG-1992 (first entry)

XX DE Der p I allergen.

XX KW House dust mite; desensitisation therapy; sensitivity.

XX OS Dermatophagoides pteronyssinus.

XX PH Key Location/Qualifiers

XX FT Cleavage-site 7..8 /note= "signal peptide cleavage"

XX FT Modified-site 68..70 /note= "N-glycosylation site"

XX PN WO9204445-A.

XX PD 19-MAR-1992.

XX PF 10-SEP-1991; 91WO-AU000417.

XX PR 11-SEP-1990; 90US-00580655.

XX PA (WAUR-) WEST AUST RES INST.

XX PI Thomas WR, Chua KY;

XX DR WPI; 1992-114353/14.

XX DR N-PSDB; AAQ23373.

XX PT DNA encoding allergens of house dust mite Dermatophagoides - and isolated protein allergens and peptide(s) useful in treating and detecting sensitivity to mites, esp. D. farinae.

XX PS Claim 14; Fig 7; 53pp; English.

XX CC



CC The protein sequence was deduced from the cDNA sequence obtd. by  
 CC screening a dermatophagoides cDNA library with two probes comprising the  
 CC Der p 1 cDNA BamHI fragments 1-348 and 349-857. The protein allergen is  
 CC used as a diagnostic reagent in detecting and treating sensitivity to  
 CC house dust mites and in desensitisation therapy. See also AAR22431.2.  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHNSQSLDLAEQ 60  
 DB 24 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHNSQSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGINYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGINYCQIY 143  
 QY 121 PPNANKIREALACTSATAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 DB 144 PPNANKIREALACTSATAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
 QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVWIL 222  
 DB 204 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVWIL 245

RESULT 4  
 AAR39359  
 ID AAR39359 standard; protein; 245 AA.  
 XX  
 AC AAR39359;  
 DT 25-MAR-2003 (revised)  
 DT 12-AUG-1993 (first entry)  
 XX  
 DE Dermatophagoides pteronyssinus Dep p1.  
 XX  
 KW T cell epitope; house dust mite; allergy; soluble; allergen.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 PN WO9308279-A1-  
 XX  
 PD 29-APR-1993.  
 XX  
 PF 15-OCT-1992; 92WO-US008637.  
 XX  
 PR 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00861396.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
 XX  
 DR WPI; 1993-152472/18.  
 DR N-PSDB; AAQ41211.  
 XX  
 PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 PT and treatment of sensitivity to house dust mite.  
 XX  
 PS Disclosure; Page 47; 176pp; English.

CC The cDNA encoding Dermatophagoides pteronyssinus group I allergen Dep p1  
 CC was obtd. in plasmid form as subclone from lambda gt11 (Chua et al., J.  
 CC Exp. Med. 167: 175-182 (1986)). The cDNA was used to express a Dep p1  
 CC allergen which was capable of raising an immune response when contacted  
 CC with anti-Dep p1 antibodies. The protein was used to design a series of  
 CC overlapping peptides synthesised by standard techniques to cover the  
 CC whole Dermatophagoides pteronyssinus Dep p1 sequence. The T cell epitopes

CC of the protein were mapped by detection of the peptide's ability to  
 CC stimulate T cell activity. The peptides may be used for diagnosis and  
 CC treatment of sensitivity to house dust mite allergens. When administered  
 CC to house dust mite sensitive individuals, the peptides are capable of  
 CC modifying the allergic response to the allergens. The peptides may be  
 CC modified for e.g. increasing solubility, enhancing therapeutic or  
 CC preventive efficacy or stability. See also AAR34586-700, AAR36398-490 and  
 CC AAR39360-2. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHNSQSLDLAEQ 60  
 DB 24 TNACSGNAPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHNSQSLDLAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGINYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIOHNGVVOESYRYVAREQSCRRPNAQRFGINYCQIY 143  
 QY 121 PPNANKIREALACTSATAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 180  
 DB 144 PPNANKIREALACTSATAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHAVNIVGYSNA 203  
 QY 181 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVWIL 222  
 DB 204 QGVYDWIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVWIL 245

RESULT 5  
 AAR51727  
 ID AAR51727 standard; protein; 245 AA.  
 XX  
 AC AAR51727;  
 DT 01-FEB-1995 (first entry)  
 XX  
 DE Der p I.  
 XX  
 KW Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
 KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;  
 KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
 KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
 KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
 KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
 KW T cell subpopulations; unresponsive; immune response; tolerance.  
 XX  
 OS Dermatophagoides pteronyssinus.

XX  
 PN ZA9302677-A.  
 XX  
 PD 26-JAN-1994.  
 XX  
 PF 16-APR-1993; 93ZA-00002677.  
 XX  
 PR 16-APR-1993; 93ZA-00002677.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX  
 DR WPI; 1994-126807/15.  
 DR N-PSDB; AAQ62308.  
 XX  
 PT Isolated and/or modified peptides comprising T-cell epitopes - of major  
 PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
 PT sensitivity to house dust mites.  
 XX  
 PS Disclosure; Page 48-49; 154pp; English.

CC This sequence represents the group I protein allergen from the house dust  
 CC mite *D. pteronyssinus*, Der p I. The Der p I protein shows high homology  
 CC to the group I protein allergen derived from *D. farinae*, Der f I, having  
 CC an identity of 81%. Fragments of these proteins, and the corresponding  
 CC group II allergens, Der p II and Der f II, (see AAR51731-841) represent  
 CC - cell epitopes. Fusion peptides may be produced which comprise at least  
 CC two or three antigenic fragments. Each region of these peptides may be  
 CC derived from the same, or different, mite allergens. The antigenic  
 CC fragments may be altered by substitution, deletion or addition to enhance  
 CC their antigenicity. These peptides may be produced by chemical synthesis,  
 CC chemical cleavage of the protein allergen or by recombinant techniques.  
 CC These peptides, when administered to a house dust mite sensitive  
 CC individual, are capable of modifying the allergic response of the  
 CC individual to the allergen. These peptides do not bind to immuno- globulin  
 CC E (IgE), or bind IgE to a lesser extent than the full length protein  
 CC allergen. This reduces the major complications of standard immunotherapy,  
 CC which are IgE-mediated responses such as anaphylaxis. Exposure of mite  
 CC allergic patients to these peptides may tolerate or anergise appropriate  
 CC T cell subpopulations such that they become unresponsive to mite  
 CC allergens and do not participate in mounting an immune response upon  
 CC exposure. Administration of the peptide may also modify the lymphokine  
 CC secretion profile as compared with exposure to the naturally occurring  
 CC mite protein allergen

XX SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACISNGNPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 24 TNACISNGNPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 143

QY 121 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYNSA 180  
 DB 144 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYNSA 203

QY 181 QGVDTYWIVRNSWDTNWGDCNGYGYFAANIDLMWIEEYPYVIL 222  
 DB 204 QGVDTYWIVRNSWDTNWGDCNGYGYFAANIDLMWIEEYPYVIL 245

RESULT 6

AAR47063  
 ID AAR47063 standard; protein; 245 AA.

XX AC AAR47063;

XX 25-MAR-2003 (revised)

DT 17-OCT-1994 (first entry)

XX Protein allergen of Derp I.

DE Der p I; House Dust Mite Allergen.

XX Dermatophagoides pteronyssinus.

XX Key Location/Qualifiers

FT Protein 1..222

FT Cleavage-site 9..111

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

FT Cleavage-site 11..23

XX 17-MAR-1994.  
 XX 10-SEP-1993; 93WO-US008518.  
 XX 10-SEP-1992; 92US-00945288.  
 XX (IMMU-) IMMUNOLOGIC PHARM CORP.

XX Thomas WR, Chua K;

XX WPI; 1994-101195/12.

XX N-PSDB; AAQ58665.

XX New protein allergens of house dust mite - used for diagnosing and  
 XX treating sensitivity in an individual to house dust mite allergens.

XX Example 1; Fig 1; 98pp; English.

XX AAR47063 is a preproenzyme Der p I. The amino acid sequence preceding the  
 XX mature protein sequence contains cleavage sites for the pre- and  
 XX proenzyme forms, with residues 1-11 corresponding to a partial signal  
 XX peptide sequence. The mature protein can be used to detect sensitivity in  
 XX an individual to house dust mite and to reduce the sensitivity of the  
 XX individual. (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNACISNGNPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

DB 24 TNACISNGNPAEIDLRQMTVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 83

QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 120

DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNAQRFGISNYCQIY 143

QY 121 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYNSA 180

DB 144 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYNSA 203

QY 181 QGVDTYWIVRNSWDTNWGDCNGYGYFAANIDLMWIEEYPYVIL 222

DB 204 QGVDTYWIVRNSWDTNWGDCNGYGYFAANIDLMWIEEYPYVIL 245

RESULT 7

AAR66545

ID AAR66545 standard; protein; 245 AA.

XX AC AAR66545;

XX 25-MAR-2003 (revised)

DT 31-JUL-1995 (first entry)

XX Der p I allergen.

XX Der p I; allergen; antigen; immunological tolerance; peptide;

XX cryptic epitope.

XX Dermatophagoides pteronyssinus.

XX WO9427634-A1.

XX 08-DEC-1994.

XX 01-JUN-1994; 94WO-AU000292.

XX 02-JUN-1993; 93US-00072832.

XX

PA (CHIL-) INST CHILD HEALTH RES.  
 XX Thomas WR;  
 PI  
 XX  
 XX WPI; 1995-022457/03.  
 DR N-PSDB; AAQ79618.  
 XX  
 XX Inducing immunological tolerance with cryptic peptide - esp from allergen  
 PT or auto:antigen, esp admin orally.  
 XX  
 XX Disclosure; Page 22-23; 38pp; English.  
 XX  
 XX Immunological tolerance to a protein antigen may be induced by admin. of  
 CC a compsn. contg. at least one cryptic peptide derived from that antigen.  
 CC Bio mice were immunised with various Der p I peptides and then response  
 CC to Der p I and the peptides was measured in the presence of spleen  
 CC adherent cells. Peptides contg. amino acids 120-143 and 144-169 of Der p  
 CC I were positive i.e., they contained cryptic epitopes. The cDNA fragment  
 CC encoding amino acids 131-187 of Der p I was cloned into pGEX and  
 CC expressed in bacteria as a fusion protein with glutathione-S-transferase.  
 CC Mice were given orally 3 mg of this protein on 3 consecutive days, then 7  
 CC days later immunised subcutaneously with native Der p I in complete  
 CC Freund's adjuvant. Seven days later lymph nodes, stimulated with protein  
 CC or synthetic peptides, were assayed for lymphokines (IL-2). Mice given  
 CC the cryptic epitope had a much weaker response than those treated only  
 CC with buffer. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 245 AA;  
 SQ

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127; Indels 0; Gaps 0;  
 Matches 222; Conservative 0; Mismatches 0;  
 Qy 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDAEQ 60  
 Db 24 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDAEQ 83  
 Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 120  
 Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 143  
 Qy 121 PNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
 Db 144 PNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 203  
 Qy 181 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
 Db 204 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 8  
 AAW71908  
 ID AAW71908 standard; protein; 245 AA.

XX AAW71908;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 16-DEC-1998 (first entry)  
 XX  
 XX Dermatophagoides Der p I.  
 XX genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
 KW Der p II; Der f I; Der f II; house dust mite allergy.  
 XX  
 XX Dermatophagoides.  
 XX US5820862-A.  
 XX  
 XX 13-OCT-1998.  
 PD  
 XX 07-JUN-1995; 95US-00482142.  
 PF  
 XX

PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
 PI Shaked Z, Rogers BL;  
 XX WPI; 1998-567590/48.  
 DR N-PSDB; AAV61384.  
 XX  
 XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.  
 FT  
 XX Disclosure; Col 69-72; 155pp; English.  
 XX  
 XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen selected  
 CC from the allergens Der p I, Der p II, Der f I, or Der f II. The invention  
 CC also describes modified peptides having similar or enhanced therapeutic  
 CC properties as the corresponding, naturally occurring allergen, but having  
 CC reduced side effects. The present sequence represents Der p I from the  
 CC present invention. (Updated on 25-MAR-2003 to correct PR field.) (Updated  
 CC on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 245 AA;  
 SQ

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127; Indels 0; Gaps 0;  
 Matches 222; Conservative 0; Mismatches 0;  
 Qy 1 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDAEQ 60  
 Db 24 TNACSGNAPAEIDLQRMRTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSILDAEQ 83  
 Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 120  
 Db 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 143  
 Qy 121 PNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 180  
 Db 144 PNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIIQRDNGYQPNYHVNIVGYSNA 203  
 Qy 181 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
 Db 204 QGVYWIWRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 245

RESULT 9  
 AAY50356  
 ID AAY50356 standard; protein; 245 AA.

XX AAY50356;  
 AC  
 XX  
 XX 25-JAN-2000 (first entry)  
 DT  
 XX Dermatophagoides sp. allergen Der p I protein fragment.  
 XX Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der p I.  
 XX  
 XX Dermatophagoides sp.  
 OS  
 XX US5968526-A.  
 XX  
 XX 19-OCT-1999.  
 PD  
 XX 07-JUN-1995; 95US-00478572.  
 PF  
 XX

PR 14-APR-1994; 94US-00227772.  
 PR 12-APR-1995; 95WO-US004481.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;  
 PI PI Evans S, Kuo M;  
 XX  
 DR WPI; 1999-590385/50.  
 DR N-PSDB; AAZ23906.  
 XX

Screening individuals for allergic reactions to T cell epitopes of major allergens from house dust mites.

Disclosure; Col 73-74; 158pp; English.

XX This invention describes a novel method (I) for detecting whether an individual is sensitive to Dermatophagoides (house dust mites). The method involves detecting sensitivity to house dust mites in patients, comprising combining a blood sample from the individual with 1 or more isolated T cell epitopes of the protein allergens I and II (DP I) and (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes with varying, defined amino acids sequences (given in the specification) may be used in (I). The sample and allergens are combined under conditions appropriate for the binding of blood components with the polypeptides. The extent of binding is then indicative of the sensitivity of the patient to house dust mites. (I) may be used to screen individuals for sensitivity to Dermatophagoides (house dust mites). The house dust mite is a major cause of a variety of allergic disorders such as asthma, rhinitis and ectopic dermatitis. This sequence represents the house dust mite allergen Der p I

XX Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLAEQ 60  
 DB 24 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLAEQ 83  
 QY 61 ELVDCASQCHGCHGTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
 DB 84 ELVDCASQCHGCHGTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 143  
 QY 121 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 DB 144 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 203  
 QY 181 QGVYDWIVRNSWNTNWDNGNGYGYFAANIDLMWIEEYPYVIL 222  
 DB 204 QGVYDWIVRNSWNTNWDNGNGYGYFAANIDLMWIEEYPYVIL 245

RESULT 10

ID AAU18959 standard; protein; 245 AA.

XX AAU18959;

DT 04-DEC-2001 (first entry)

DE House dust mite allergen Der p I.

XX House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.

OS Dermatophagoides pteronyssinus.

XX US6268491-B1.

XX 31-JUL-2001.  
 XX 07-JUN-1995; 95US-00484296.  
 XX 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.

XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI PI Evans S, Shaked Z;  
 XX

DR WPI; 2001-549074/61.  
 DR N-PSDB; AAS30721.

XX Peptides comprising T cell groups of the major allergens from Dermatophagoides (house dust mites), useful for treating house dust mite allergy in humans, and for diagnosing sensitivity to house dust mite protein allergens.

XX Example 1; Col 73; 158pp; English.

XX The invention relates to an isolated peptide of the major protein allergens of the genus Dermatophagoides, which comprises at least one T cell group of a protein allergen from Der p (DP) I, DP II, Der f (DP) I or DP II. The isolated peptide comprises at least two regions, each region comprising at least one T cell group of a protein allergen of the genus Dermatophagoides. The regions are derived from the same or different protein allergens of the genus Dermatophagoides. The peptides are useful for treating house dust mite allergy in humans. The peptides are also useful for detecting or diagnosing sensitivity to house dust mite protein allergens. The present peptides have similar or enhanced therapeutic properties as the naturally-occurring allergen, but have reduced side effects, and increased solubility and stability. The present sequence represents an allergenic protein from Dermatophagoides from which the T-cell epitope containing peptides are derived

XX Sequence 245 AA;

Query Match 100.0%; Score 1206; DB 4; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLAEQ 60  
 DB 24 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLAEQ 83  
 QY 61 ELVDCASQCHGCHGTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120  
 DB 84 ELVDCASQCHGCHGTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 143  
 QY 121 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 DB 144 PPNAKIREALAQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 203  
 QY 181 QGVYDWIVRNSWNTNWDNGNGYGYFAANIDLMWIEEYPYVIL 222  
 DB 204 QGVYDWIVRNSWNTNWDNGNGYGYFAANIDLMWIEEYPYVIL 245

RESULT 11

ID ABG71812 standard; protein; 245 AA.

XX ABG71812;

XX 10-JAN-2003 (first entry)

DE European house dust mite Der p I protein.

XX Dermatophagoides farinae protein allergen; Der f I; Der f II; allergy;  
 KW allergic response; mite allergen; house dust mite; T cell epitope;  
 KW B cell epitope; anti-allergic; desensitisation therapy;  
 KW modifies allergic response of house dust mite-sensitive individual;  
 KW modify B-cell and/or T-cell response to Der f I and Der f II;  
 KW European house dust mite; Der p I.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 XX US6423837-B1.  
 XX  
 PD 23-JUL-2002.  
 XX  
 XX 20-APR-1999; 99US-00295188.  
 XX  
 PR 13-FEB-1990; 90US-00458642.  
 PR 11-SEP-1990; 90US-00580655.  
 PR 16-AUG-1993; 93US-00107332.  
 PR 29-DEC-1993; 93US-00175071.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Thomas WR, Chua K;  
 XX  
 XX WPI; 2002-672946/72.  
 XX N-PSDB; ABS56342.  
 XX  
 PT Novel isolated DNA encoding peptide from Dermatophagoides farinae protein  
 PT allergen, designated Der f I and Der f II, useful for treating and  
 PT preventing allergic responses to mite allergens, by desensitization  
 PT therapy.  
 XX  
 XX Disclosure; Fig 7; 22pp; English.  
 PS  
 XX The present invention relates to a new DNA encoding a peptide from  
 CC Dermatophagoides farinae protein allergen, designated Der f I or Der f  
 CC II, that comprises at least one epitope of the protein allergen. The  
 CC invention is useful for producing a peptide from Der f I or Der f II  
 CC protein allergen. The invention is also useful for diagnosing, treating  
 CC and preventing allergic responses to mite allergens, particularly, the  
 CC mites D. farinae. The invention is also useful as a probe for identifying  
 CC additional nucleotide sequences coding for mite allergens having amino  
 CC acid sequences similar to Der f I or Der f II. The peptide is useful for  
 CC detecting sensitivity in an individual to house dust mites and can be  
 CC used to treat sensitivity (reduce sensitivity or desensitise) in an  
 CC individual, to whom therapeutically effective quantities of D. farinae  
 CC peptide is administered. The peptides when administered to sensitive  
 CC individual modify the individual's allergic response to Der f I or Der f  
 CC II. The peptides are also useful as purified allergens useful in the  
 CC standardisation of allergen extracts or preparations which can be used as  
 CC reagents for diagnosis and treatment of allergy to house dust mites. The  
 CC epitopic peptides are useful as diagnostic reagents for determining  
 CC reactivity to the mite species. The peptides are also useful for  
 CC identifying or defining T cell epitopes and/or B cell epitopes which are  
 CC of importance in allergic responses to D. farinae allergens and to  
 CC elucidate the mediators or mechanisms of by which these reactions occur.  
 CC The present amino acid sequence represents the European house dust mite  
 CC Der p I protein as described in the invention  
 XX  
 XX Sequence 245 AA;  
 SQ  
 Query Match 100.0%; Score 1206; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127; Indels 0; Gaps 0;  
 Matches 222; Conservative 0; Mismatches 0;  
 QY 1 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 24 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 83  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 120  
 DB 84 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 143

QY 121 PPANKIREALQTHSAIAVIIGIKDAPFRHYDGTIQRDNGYQPNYHVNIVGYSNA 180  
 DB 144 PPANKIREALQTHSAIAVIIGIKDAPFRHYDGTIQRDNGYQPNYHVNIVGYSNA 203  
 QY 181 QGVYDIIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 222  
 DB 204 QGVYDIIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPYVIL 245  
 RESULT 12  
 AAR49920  
 ID AAR49920 standard; protein; 320 AA.  
 XX  
 AC AAR49920;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 17-OCT-1994 (first entry)  
 XX  
 DE Protein allergen of Der pI.  
 XX  
 KW Der pI; House Dust Mite Allergen.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..322  
 FT /label= Der pI preproenzyme  
 FT Cleavage-site 83..85  
 FT /label= Cleavage site  
 FT /note= "proenzyme remains"  
 FT Cleavage-site 85..97  
 FT /label= cleavage site  
 FT /note= "cleavage between pro-Der pI and pre-Der pI"  
 XX  
 XX W09405790-A1.  
 EN 17-MAR-1994.  
 PD 10-SEP-1993; 93WO-US008518.  
 XX  
 PR 10-SEP-1992; 92US-00945288.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA Thomas WR, Chua K;  
 PI WPI; 1994-101195/12.  
 DR N-PSDB; AAQ38669.  
 XX  
 PT New protein allergens of house dust mite - used for diagnosing and  
 PT treating sensitivity in an individual to house dust mite allergens.  
 XX  
 XX Example 1; Fig 21; 98pp; English.  
 CC AAR49920 is a preproenzyme Der pI. The amino acid sequence preceding the  
 CC mature protein sequence contains cleavage sites for the pre- and  
 CC proenzyme forms, with residues 1-97 corresponding to a partial signal  
 CC peptide sequence. The mature protein can be used to detect sensitivity in  
 CC an individual to house dust mite and to reduce the sensitivity of the  
 CC individual. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 320 AA;  
 SQ  
 Query Match 100.0%; Score 1206; DB 2; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-127;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60  
 DB 99 TNACISNGNAPAEIDLRQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 158  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRPNQRFGISNYCQIY 120

Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVAREQSCRRPNAQRFGISNYCQIY 218  
 QY 121 PPNANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 180  
 Db 219 PPNANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 278  
 QY 181 QGVDTWIVRNSWDTNWGDNGGYGFAANIDLMWIEEYPYVIL 222  
 Db 279 QGVDTWIVRNSWDTNWGDNGGYGFAANIDLMWIEEYPYVIL 320

RESULT 13  
 AAU07748  
 ID AAU07748 standard; protein; 222 AA.  
 XX  
 AC AAU07748;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE House dust mite allergenic protein Der p I variant d.  
 XX  
 KW House dust mite; allergenic protein; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope; polymorphic variant.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 136  
 FT /note= "Wild-type Ser substituted by Thr"  
 XX  
 FN US6268491-B1.  
 XX  
 PD 31-JUL-2001.  
 XX  
 PF 07-JUN-1995; 95US-00484296.  
 XX  
 PR 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-0045307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 XX  
 DR WPI; 2001-549074/61.  
 XX  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX  
 XX Disclosure; Fig 22; 158pp; English.  
 XX  
 CC The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic protein from Dermatophagoides from  
 CC which the T-cell epitope containing peptides are derived, a polymorphic  
 CC variant of Der p I. Note: The present sequence is not shown in the

CC specification but is derived from the Der p I sequence shown in figure 22  
 XX  
 SQ Sequence 222 AA;  
 Query Match 99.8%; Score 1203; DB 4; Length 222;  
 Best Local Similarity 99.5%; Pred. No. 5.5e-127;  
 Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACISNGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAAATESAYLAHRNQLDLAEQ 60  
 Db 1 TNACISNGNAPAEIDLQRMETVTPIRMGGCGSCWAFSGVAAATESAYLAHRNQLDLAEQ 60  
 QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVAREQSCRRPNAQRFGISNYCQIY 120  
 Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRVAREQSCRRPNAQRFGISNYCQIY 120  
 QY 121 PPNANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 180  
 Db 121 PPNANKIREALAOHTSAIAVIIGIKDLDAFRHYDGRITIIORDNGYQPNYHNAVIVGYSNA 180  
 QY 181 QGVDTWIVRNSWDTNWGDNGGYGFAANIDLMWIEEYPYVIL 222  
 Db 181 QGVDTWIVRNSWDTNWGDNGGYGFAANIDLMWIEEYPYVIL 222

RESULT 14  
 AAB98347  
 ID AAB98347 standard; protein; 222 AA.  
 XX  
 AC AAB98347;  
 XX  
 DT 21-AUG-2001 (first entry)  
 XX  
 DE D. pteronyssinus Der p I protein SEQ ID NO:82.  
 XX  
 KW Mite group I protein; methylotropic yeast; Escherichia coli; allergy;  
 KW recombinant mite group I protein; allergic response; antiallergic;  
 KW infectious disease; allergic disease.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 WO200129078-A2.  
 XX  
 PD 26-APR-2001.  
 XX  
 PF 12-OCT-2000; 2000WO-US028204.  
 XX  
 PR 15-OCT-1999; 99US-0159841P.  
 XX  
 PA (HESK-) HESKA CORP.  
 XX  
 PI Best EA, McDermott MJ;  
 XX  
 XX WPI; 2001-308475/32.  
 DR N-P8DB; AAH22385.  
 XX  
 PT Producing recombinant mite Group I protein for treating allergies,  
 PT involves culturing a methylotropic yeast microorganism or Escherichia  
 PT coli transformed with nucleic acid molecule, and recovering the protein.  
 XX  
 PS Claim 12; Page 145; 154pp; English.  
 CC The present invention describes a method for the production of a  
 CC recombinant mite Group I protein (I). The method comprises culturing a  
 CC methylotropic yeast microorganism transformed with a nucleic acid  
 CC molecule (II) encoding (I), and recovering (I), or culturing Escherichia  
 CC coli transformed with (II) under conditions in which (I) forms an  
 CC inclusion body in E. coli, isolating the inclusion body, and recovering  
 CC (I). Also described is a method for detecting mite allergy in an animal  
 CC comprising: (a) contacting (I) with a putative IGE-containing substance  
 CC to form a complex between (I) and IGE; and (b) determining the presence  
 CC of IGE reactive with (I) by detecting the complex, where the presence of  
 CC reactive IGE is indicative of mite allergy in the animal. (I) is useful

CC for detecting mite allergy in an animal, or in a composition to reduce  
 CC allergic response to a mite Group 1 protein in a mite allergic animal.  
 CC (I) is also useful in a composition for treating or preventing allergic,  
 CC infectious or other diseases. AAH22326 to AAH22394 and AAB98326 to  
 CC AAB98349 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 222 AA;

Query Match 99.5%; Score 1200; DB 4; Length 222;  
 Best Local Similarity 99.5%; Pred. No. 1.2e-126;  
 Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLQRMETVPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60  
 DB 1 TNACSGNAPAEIDLQRMETVPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60  
 QY 61 ELVDCAHQCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNQAFGINSYQIY 120  
 DB 61 ELVDCAHQCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNQAFGINSYQIY 120  
 QY 121 PPNAKIREALAOHTSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
 DB 121 PPNAKIREALAOHTSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
 QY 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMMEIEYPYVWIL 222  
 DB 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMMEIEYPYVWIL 222

RESULT 15  
 AAB98346  
 ID AAB98346 standard; protein; 302 AA.

AC AAB98346;  
 XX  
 XX 21-AUG-2001 (first entry)  
 XX  
 DE D. pteronyssinus Der p 1 protein SEQ ID NO:77.  
 XX  
 XX Mite group 1 protein; methylophilic yeast; Escherichia coli; allergy;  
 XX recombinant mite group 1 protein; allergic response; antiallergic;  
 XX infectious disease; allergic disease.  
 XX

Dermatophagoides pteronyssinus.

WO200129078-A2.

26-APR-2001.

12-OCT-2000; 2000WO-US028204.

15-OCT-1999; 99US-0159841P.

(HESK-) HESKA CORP.

Best EA, McDermott MJ;

WPI; 2001-308475/32.

N-PSDB; AAH22381.

PT Producing recombinant mite Group 1 protein for treating allergies,  
 PT involves culturing a methylophilic yeast microorganism or Escherichia  
 PT coli transformed with nucleic acid molecule, and recovering the protein.

Claim 12; Page 141-142; 154pp; English.

CC The present invention describes a method for the production of a  
 CC recombinant mite Group 1 protein (I). The method comprises culturing a  
 CC methylophilic yeast microorganism transformed with a nucleic acid  
 CC molecule (II) encoding (i), and recovering (i), or culturing Escherichia  
 CC coli transformed with (ii) under conditions in which (i) forms an  
 CC inclusion body in E. coli, isolating the inclusion body, and recovering

CC (I). Also described is a method for detecting mite allergy in an animal  
 CC comprising: (a) contacting (i) with a putative IgE-containing substance  
 CC to form a complex between (i) and IgE; and (b) determining the presence  
 CC of IgE reactive with (i) by detecting the complex, where the presence of  
 CC reactive IgE is indicative of mite allergy in the animal. (I) is useful  
 CC for detecting mite allergy in an animal, or in a composition to reduce  
 CC allergic response to a mite Group 1 protein in a mite allergic animal.  
 CC (I) is also useful in a composition for treating or preventing allergic,  
 CC infectious or other diseases. AAH22326 to AAH22394 and AAB98326 to  
 CC AAB98349 represent sequences used in the exemplification of the present  
 CC invention  
 XX

SQ Sequence 302 AA;

Query Match 99.5%; Score 1200; DB 4; Length 302;  
 Best Local Similarity 99.5%; Pred. No. 1.9e-126;  
 Matches 221; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNACSGNAPAEIDLQRMETVPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLAEQ 60  
 DB 81 TNACSGNAPAEIDLQRMETVPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLAEQ 140  
 QY 61 ELVDCAHQCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNQAFGINSYQIY 120  
 DB 141 ELVDCAHQCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRPNQAFGINSYQIY 200  
 QY 121 PPNAKIREALAOHTSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 180  
 DB 201 PPNAKIREALAOHTSAIAVIIGIKDLDAFRHYDGTIIQRDNGYQPNYHVNIVGYSNA 260  
 QY 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMMEIEYPYVWIL 222  
 DB 261 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMMEIEYPYVWIL 302

Search completed: October 27, 2004, 17:43:56  
 Job time : 309.024 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:50:34 ; Search time 229.96 Seconds  
(without alignments)

312.993 Million cell updates/sec

Title: US-09-867-159A-2

Perfect score: 1206

Sequence: 1 TNACSGNAPAEIDLQRMR.....YFAANIDLMIEEYFYVIL 222

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1206	100.0	222	10	US-09-867-159A-2
2	1196	99.2	222	14	US-10-001-245-88
3	1196	99.2	320	9	US-09-877-160-2
4	1196	99.2	320	10	US-09-847-208-79
5	1196	99.2	320	14	US-10-001-245-179
6	1182	96.4	222	14	US-10-001-245-26
7	1161	96.3	222	14	US-10-001-245-18
8	1161	96.3	222	14	US-10-001-245-20
9	1161	96.3	222	14	US-10-001-245-34
10	1159	96.1	222	14	US-10-001-245-22
11	1157	95.9	222	14	US-10-001-245-14
12	1157	95.9	222	14	US-10-001-245-16
13	1143	94.8	222	14	US-10-001-245-30

```
14 1142 94.7 222 14 US-10-001-245-28 Sequence 28, Appl
15 1133 93.9 222 14 US-10-001-245-34 Sequence 34, Appl
16 1132 93.9 222 14 US-10-001-245-32 Sequence 32, Appl
17 1040.5 86.3 327 14 US-10-001-245-182 Sequence 182, App
18 1037.5 86.0 246 14 US-10-001-245-181 Sequence 181, App
19 1037.5 86.0 321 14 US-10-001-245-180 Sequence 180, App
20 997.5 82.7 321 10 US-09-847-208-73 Sequence 73, Appl
21 997.5 82.7 321 14 US-10-001-245-183 Sequence 183, App
22 982.5 81.5 211 10 US-09-847-208-95 Sequence 95, Appl
23 982.5 81.5 211 14 US-10-001-245-184 Sequence 184, App
24 944.5 78.3 210 14 US-10-001-245-185 Sequence 185, App
25 332.5 27.6 357 16 US-10-437-963-135411 Sequence 135411,
26 332.5 27.6 696 14 US-10-259-165-184 Sequence 184, App
27 331 27.4 363 15 US-10-425-114-48978 Sequence 48978, A
28 324.5 26.9 249 15 US-10-425-114-44401 Sequence 44401, A
29 324.5 26.9 381 15 US-10-425-114-43777 Sequence 43777, A
30 323.5 26.8 329 9 US-09-953-956-2 Sequence 2, Appli
31 323.5 26.8 329 13 US-10-114-464-2 Sequence 2, Appli
32 323.5 26.8 329 14 US-10-318-584-2 Sequence 2, Appli
33 323.5 26.8 361 9 US-09-925-300-1693 Sequence 1693, Ap
34 319.5 26.5 373 14 US-10-219-220-285 Sequence 285, App
35 319.5 26.5 396 14 US-10-219-220-153 Sequence 153, App
36 319 26.5 349 16 US-10-437-963-104361 Sequence 104361,
37 318 26.4 436 14 US-10-259-165-108 Sequence 108, App
38 318 26.4 436 14 US-10-259-165-442 Sequence 442, App
39 318 26.4 447 15 US-10-425-114-68471 Sequence 68471, A
40 318 26.4 458 16 US-10-437-963-155525 Sequence 155525,
41 318 26.4 461 15 US-10-425-114-58026 Sequence 58026, A
42 317.5 26.3 329 9 US-09-953-956-7 Sequence 7, Appli
43 317.5 26.3 329 13 US-10-114-464-7 Sequence 7, Appli
44 317.5 26.3 329 14 US-10-318-584-3 Sequence 3, Appli
45 317 26.3 500 14 US-10-219-220-149 Sequence 149, App
```

#### ALIGNMENTS

```
US-09-867-159A-2
; Sequence 2, Application US/09867159A
; Publication No. US20030104013A1
; GENERAL INFORMATION:
; APPLICANT: ANTIALLIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one
; TITLE OF INVENTION: and at least one anti-histamine compound
; FILE REFERENCE: B112812US-antiallis
; CURRENT APPLICATION NUMBER: US/09/867,159A
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: FR01/04370
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: FR01/05929
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(222)
; OTHER INFORMATION: Peptide sequence from cystine protease.
```

```
US-09-867-159A-2
Query Match 100.0%; Score 1206; DB 10; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.8e+118; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;
Qy 1 TNACSGNAPAEIDLQRMTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSDLAEQ 60
Db 1 TNACSGNAPAEIDLQRMTVTPIRMQCGCGSCWAFSGVAATESAYLAHRNOSDLAEQ 60
Qy 61 ELVDCAQSCHSGDTPRGIEYTHNGVTVQESYRVYVAREQSCRRPNAQRFPGISNYCQIV 120
```

```

; LENGTH: 320
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-09-877-160-2

Query Match      99.2%; Score 1196; DB 14; Length 222;
Best Local Similarity 99.1%; Pred. No. 3.4e-117;
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60
Db 99 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 158

Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 218

Qy 121 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
Db 219 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278

Qy 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 222
Db 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 320

```

## RESULT 2

```

US-10-001-245-88
; Sequence 88 Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-88

```

```

Query Match      99.2%; Score 1196; DB 14; Length 222;
Best Local Similarity 99.1%; Pred. No. 2.1e-117;
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60
Db 1 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60

Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120
Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120

Qy 121 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
Db 121 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180

Qy 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 222
Db 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 222

```

## RESULT 3

```

US-09-877-160-2
; Sequence 2, Application US/09877160
; Publication No. US20020197269A1
; GENERAL INFORMATION:
; APPLICANT: Ching-Hsiao, Hsu
; APPLICANT: Cheng, Winston T. K.
; TITLE OF INVENTION: ALLERGEN-CONTAINING MILK FOR ALLERGY
; FILE REFERENCE: 12774-003001
; CURRENT APPLICATION NUMBER: US/09/877,160
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Vers:0n 4.0
; SEQ ID NO 2

```

```

; LENGTH: 320
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-09-877-160-2

Query Match      99.2%; Score 1196; DB 9; Length 320;
Best Local Similarity 99.1%; Pred. No. 3.4e-117;
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60
Db 99 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 158

Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 218

Qy 121 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
Db 219 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278

Qy 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 222
Db 279 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 320

RESULT 4
US-09-847-208-79
; Sequence 79, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhu, Daocheng
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus (House-dust mite)
US-09-847-208-79

Query Match      99.2%; Score 1196; DB 10; Length 320;
Best Local Similarity 99.1%; Pred. No. 3.4e-117;
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 60
Db 99 TNACISNGNAPAEIDLQMETVTPIRMQGGCGSWAFSGVAATESAYLAHRNQSLLDAEQ 158

Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 120
Db 159 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAORFGISNYCQIY 218

Qy 121 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180
Db 219 PPVANKIREALAOHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 278

Qy 181 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 222
Db 279 QGVDTWIVRNSWDTNWGDNGYGYFAANIDLMWIEEYPPVVIL 320

RESULT 5
US-10-001-245-179
; Sequence 179, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:

```

; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 179  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-179

Query Match 99.2%; Score 1196; DB 14; Length 320;  
Best Local Similarity 99.1%; Pred. No. 3.4e-117;  
Matches 220; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMETVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Db 99 TNACSGNAPAEIDLQRMETVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 159

QY 61 ELVDCASQCHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 159 ELVDCASQCHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 218

QY 121 PPVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQDNGYQPNYHAVNIVGYSNA 180  
Db 219 PPVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQDNGYQPNYHAVNIVGYSNA 278

QY 181 QGVYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222  
Db 279 QGVYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 320

## RESULT 6

US-10-001-245-26  
; Sequence 26, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 26  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-26

Query Match 96.4%; Score 1162; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 7.9e-114;  
Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMETVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Db 1 TNACSGNAPAEIDLQRMETVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

QY 61 ELVDCASQCHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 61 ELVDCASQCHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120

QY 121 PPVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQDNGYQPNYHAVNIVGYSNA 180  
Db 121 PPVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQDNGYQPNYHAVNIVGYSNA 180

QY 181 QGVYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222  
Db 181 QGVYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222

## RESULT 7

US-10-001-245-18  
; Sequence 18, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 18  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-18

Query Match 96.3%; Score 1161; DB 14; Length 222;  
Best Local Similarity 96.4%; Pred. No. 1e-113;  
Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 TNACSGNAPAEIDLQRMETVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60  
Db 1 TNACSGNAPAEIDLQRMETVPIRMQGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60

QY 61 ELVDCASQCHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120  
Db 61 ELVDCASQCHGCHGDTIPRGIEYIQHNGVQESYRYVAREQSCRRPNAQRFGISNYCQIY 120

QY 121 PPVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQDNGYQPNYHAVNIVGYSNA 180  
Db 121 PPVANKIREALAQTHSAIAVIIGIKDLDAFRHYDGTIIQDNGYQPNYHAVNIVGYSNA 180

QY 181 QGVYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222  
Db 181 QGVYWIVRNSWDTNWGDNGYGYFAANIDLMIEEYPYVIL 222

## RESULT 8

US-10-001-245-20  
; Sequence 20, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/298,170  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/249,361  
 ; PRIOR FILING DATE: 2000-11-16  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 20  
 ; LENGTH: 222  
 ; TYPE: PRT  
 ; ORGANISM: Dermatophagoides pteronyssinus  
 US-10-001-245-20

Query Match 96.3%; Score 1161; DB 14; Length 222;  
 Best Local Similarity 96.4%; Pred. No. 1e-113;  
 Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 TNACISNGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 Db 1 TNACISNGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAEQQSCRRPNAORFGISNYCOIY 120  
 Db 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAEQQSCRRPNAORFGISNYCOIY 120  
 Qy 121 PPNANKIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 Db 121 PPNVKNIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 Qy 181 QGVDTWIVRNSWDNTNWDNGYGYFAANIDLMIMIEEYFYVWIL 222  
 Db 181 QGVDTWIVRNSWDNTNWDNGYGYFAANIDLMIMIEEYFYVWIL 222

RESULT 9  
 US-10-001-245-24  
 ; Sequence 24, Application US/10001245  
 ; Publication No. US20030175312A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLM, Jens  
 ; APPLICANT: IPSEN, Henrik  
 ; APPLICANT: LARSEN, Jorgen N.  
 ; APPLICANT: SPANGFORT, Michael D.  
 ; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
 ; FILE REFERENCE: 4305/1H942-US2  
 ; CURRENT APPLICATION NUMBER: US/10/001,245  
 ; CURRENT FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/298,170  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/249,361  
 ; PRIOR FILING DATE: 2000-11-16  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 24  
 ; LENGTH: 222  
 ; TYPE: PRT  
 ; ORGANISM: Dermatophagoides pteronyssinus  
 US-10-001-245-24

Query Match 96.3%; Score 1161; DB 14; Length 222;  
 Best Local Similarity 96.4%; Pred. No. 1e-113;  
 Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 TNACISNGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 Db 1 TNACISNGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAEQQSCRRPNAORFGISNYCOIY 120  
 Db 61 ELVDCANQHGCHGDTIPRGIEYIQHNGVQESYRYVAEQQSCRRPNAORFGISNYCOIY 120  
 Qy 121 PPNANKIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 Db 121 PPNVKNIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180

Qy 181 QGVDTWIVRNSWDNTNWDNGYGYFAANIDLMIMIEEYFYVWIL 222  
 Db 181 QGVDTWIVRNSWDNTNWDNGYGYFAANIDLMIMIEEYFYVWIL 222

RESULT 10  
 US-10-001-245-22  
 ; Sequence 22, Application US/10001245  
 ; Publication No. US20030175312A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLM, Jens  
 ; APPLICANT: IPSEN, Henrik  
 ; APPLICANT: LARSEN, Jorgen N.  
 ; APPLICANT: SPANGFORT, Michael D.  
 ; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
 ; FILE REFERENCE: 4305/1H942-US2  
 ; CURRENT APPLICATION NUMBER: US/10/001,245  
 ; CURRENT FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/298,170  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/249,361  
 ; PRIOR FILING DATE: 2000-11-16  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 22  
 ; LENGTH: 222  
 ; TYPE: PRT  
 ; ORGANISM: Dermatophagoides pteronyssinus  
 US-10-001-245-22

Query Match 96.1%; Score 1159; DB 14; Length 222;  
 Best Local Similarity 96.4%; Pred. No. 1.6e-113;  
 Matches 214; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 1 TNACISNGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 Db 1 TNACISNGNAPAEIDLQMTVTPIRMQGGCGSWAFSGVAATESAYLAHRNOSLDLAEQ 60  
 Qy 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVQESYRYVAEQQSCRRPNAORFGISNYCOIY 120  
 Db 61 ELVDCANQHGCHGDTIPRGIEYIQHNGVQESYRYVAEQQSCRRPNAORFGISNYCOIY 120  
 Qy 121 PPNANKIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 Db 121 PPNVKNIREALAOQTHSAIAVIGIKDLDAFRHYDGRITIIQDNGYQPNYHVNIVGYSNA 180  
 Qy 181 QGVDTWIVRNSWDNTNWDNGYGYFAANIDLMIMIEEYFYVWIL 222  
 Db 181 QGVDTWIVRNSWDNTNWDNGYGYFAANIDLMIMIEEYFYVWIL 222

RESULT 11  
 US-10-001-245-14  
 ; Sequence 14, Application US/10001245  
 ; Publication No. US20030175312A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLM, Jens  
 ; APPLICANT: IPSEN, Henrik  
 ; APPLICANT: LARSEN, Jorgen N.  
 ; APPLICANT: SPANGFORT, Michael D.  
 ; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
 ; FILE REFERENCE: 4305/1H942-US2  
 ; CURRENT APPLICATION NUMBER: US/10/001,245  
 ; CURRENT FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/298,170  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/249,361  
 ; PRIOR FILING DATE: 2000-11-16  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 14  
 ; LENGTH: 222

```
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-14

Query Match
Best Local Similarity 95.9%; Score 1157; DB 14; Length 222;
Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
DB 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
QY 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222
DB 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222

RESULT 12
US-10-001-245-15
; Sequence 16, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-16

Query Match
Best Local Similarity 95.9%; Score 1157; DB 14; Length 222;
Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
DB 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
QY 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222
DB 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222

RESULT 13
US-10-001-245-30
; Sequence 28, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-16

Query Match
Best Local Similarity 95.9%; Score 1157; DB 14; Length 222;
Matches 214; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
DB 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
QY 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222
DB 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222

RESULT 13
US-10-001-245-30
; Sequence 28, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 28
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-28

Query Match
Best Local Similarity 94.7%; Score 1142; DB 14; Length 222;
Matches 211; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
```

```
; Sequence 30, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-30

Query Match
Best Local Similarity 94.8%; Score 1143; DB 14; Length 222;
Matches 211; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
DB 1 TNACSSINGNAPAEIDLQRMRTVTPIRMGGCGSCWAFSGVAATESAYLAHRNQSLLDAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRYVAREQSCRRPNQAFGINSYCOIY 120
QY 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
DB 121 PPNANKIREALQTHSAIAVIIGIKDLDAFRHYDGRITIQDNGYQPNYHVNIVGYGNA 180
QY 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222
DB 181 QGVYDWIVRNSWDTNWGNGYGYFAANIDLMWIEEYPYVWIL 222

RESULT 14
US-10-001-245-28
; Sequence 28, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 28
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-28

Query Match
Best Local Similarity 95.0%; Pred. No. 1e-111;
Matches 211; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSIDLAEQ 60
DB 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSIDLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAORFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAORFGISNYCQIY 120
QY 121 PPANKIREALAQTHSAIAVIGIKLDAPRHYDGTIIQRDNGYQPNYHAYNIVGYSNA 180
DB 121 PPANKIREALAQTHSAIAVIGIKLDAPRHYDGTIIQRDNGYQPNYHAYNIVGYSNA 180
QY 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPPYVIL 222
DB 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPPYVIL 222
```

## RESULT 15

```
US-10-001-245-34
; Sequence 34, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-34
```

```
Query Match 93.9%; Score 1133; DB 14; Length 222;
Best Local Similarity 94.6%; Pred. No. 8.9e-111;
Matches 210; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSIDLAEQ 60
DB 1 TNACSGNAPAEIDLQRMETVTPIRMQGGCGSCWAFSGVAATESAYLAHRNQSIDLAEQ 60
QY 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAORFGISNYCQIY 120
DB 61 ELVDCASQHGCHGDTIPRGIEYIQHNGVVOESYRVRVAREQSCRRPNAORFGISNYCQIY 120
QY 121 PPANKIREALAQTHSAIAVIGIKLDAPRHYDGTIIQRDNGYQPNYHAYNIVGYSNA 180
DB 121 PPANKIREALAQTHSAIAVIGIKLDAPRHYDGTIIQRDNGYQPNYHAYNIVGYSNA 180
QY 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPPYVIL 222
DB 181 QGVDDYWIIVRNSWDTNMGDNGYGYFAANIDLMIMEEYPPYVIL 222
```

Search completed: October 27, 2004, 18:15:42  
Job time : 230.96 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:12 ; Search time 2.66932 Seconds  
(without alignments)  
360.454 Million cell updates/sec

Title: US-09-867-159a-3

Perfect score: 61

Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	90.2	211	2 S21864	probable cysteine
2	55	90.2	245	2 JQ0337	allergen Der p 1 -
3	55	90.2	319	2 A61500	allergen Der f 1 p
4	46	75.4	76	1 S78753	hypothetical prote
5	46	75.4	76	2 AB1370	hypothetical prote
6	45	73.8	95	2 FQ0650	senescence-associat
7	45	73.8	221	2 A59041	cysteine proteinas
8	45	73.8	319	2 A59040	cysteine proteinas
9	45	73.8	319	2 T08844	cysteine proteinas
10	45	73.8	458	1 KHRZOA	oryzain (EC 3.4.22
11	45	73.8	462	2 JN0719	drought-inducible
12	45	73.8	658	2 T08153	cysteine proteinas
13	44	72.1	377	2 T12042	cysteine proteinas
14	43	70.5	224	2 T04246	hypothetical prote
15	43	70.5	343	2 JCY183	cathepsin Q (EC 3.
16	42	68.9	87	2 S50134	nitrogen fixation
17	42	68.9	94	2 S03380	major fecal allerg
18	42	68.9	300	1 D34443	nitrogen fixation
19	42	68.9	300	2 AD1988	nitrogen fixation
20	42	68.9	348	2 JN0633	caricain (EC 3.4.2
21	42	68.9	367	2 JN0634	caricain (EC 3.4.2
22	42	68.9	1364	2 T10236	xanthine dehydroge
23	41	67.2	88	2 JQ0514	cysteine-rich 9K p
24	41	67.2	88	2 S12126	cysteine-rich oute
25	41	67.2	88	2 S12125	cysteine-rich oute
26	41	67.2	88	2 E71513	probable 9kDa-cyst
27	41	67.2	88	2 D81671	cysteine rich oute
28	41	67.2	112	2 I39609	nifu protein - Ana
29	41	67.2	164	2 T24272	hypothetical prote

30 41 67.2 169 1 S18946 ultra high-sulfur  
31 41 67.2 182 2 A36886 ultra-high-sulfur  
32 41 67.2 188 2 T15851 hypothetical prote  
33 41 67.2 191 2 I46412 keratin KAP5.4 - s  
34 41 67.2 197 2 I46413 keratin KAP5.5 - s  
35 41 67.2 223 2 B38346 ultra-high-sulfur  
36 41 67.2 312 2 A23705 cysteine proteinas  
37 41 67.2 335 1 KHHOB cathepsin B (EC 3.  
38 41 67.2 339 1 KHHUB cathepsin B (EC 3.  
39 41 67.2 339 1 KHMBS cathepsin B (EC 3.  
40 41 67.2 339 1 KHRTB cathepsin B (EC 3.  
41 41 67.2 340 2 S58770 cathepsin B (EC 3.  
42 41 67.2 341 2 T29598 hypothetical prote  
43 41 67.2 344 2 S38939 probable cathepsin  
44 41 67.2 348 2 S38148 hypothetical prote  
45 41 67.2 380 2 S55923 cysteine proteinas

#### ALIGNMENTS

##### RESULT 1

S21864

probable cysteine proteinase (EC 3.4.22.-) - Euroglyphus maynei

N:Alternate names: allergen Eur m I

C:Species: Euroglyphus maynei

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S21864

R:Kent, N.A.; Hill, M.; Keen, J.N.; Holland, P.W.H.; Hart, B.

submitted to the EMBL Data Library, June 1991

A:Reference number: S21864

A:Accession: S21864

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-211 <KEN>

A:Cross-references: UNIPROT:P25780; EMBL:X60073

C:Genetics:

A:Introns: 100/3; 155/2

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

Query Match 90.2%; Score 55; DB 2; Length 211;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9

DB 27 RMQGGCGSC 35

##### RESULT 2

JQ0337

allergen Der p 1 - house-dust mite (Dermatophagoides pteronyssinus) (fragment)

C:Species: Dermatophagoides pteronyssinus

C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C:Accession: JQ0337; A27582; A31657; C27634

R:Chua, K.Y.; Stewart, G.A.; Thomas, W.R.; Simpson, R.J.; Plozza, T.M.; J

J. Exp. Med. 167, 175-182, 1988

A>Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1:

A:Reference number: JQ0337; MUID:88089411; PMID:3335830

A:Accession: JQ0337

A:Molecule type: mRNA

A:Residues: 1-245 <CHU>

A:Cross-references: UNIPROT:P08176

R:Thomas, W.R.; Stewart, G.A.; Simpson, R.J.; Chua, K.Y.; Plozza, T.M.; Dilworth, R.J.; J

Int. Arch. Allergy Appl. Immunol. 85, 127-129, 1988

A>Title: Cloning and expression of DNA coding for the major house dust mite allergen Der

A:Reference number: A27582; MUID:88114080; PMID:3276629

A:Accession: A27582

A:Molecule type: mRNA

A:Residues: 6-101 <THO>

A:Cross-references: GB:M24794; NID:g387591; PIDN:AAA28296.1; PID:g387592

R:Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.

Protein Seq. Data Anal. 2, 17-21, 1999  
 A:Title: Structural studies on the allergen Der p1 from the house dust mite Dermatophagoides pteronyssinus  
 A:Reference number: A31657; MUID:89098855; PMID:2911558  
 A:Accession: A31657  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 24-56, 'XX', 59-64; 102-118; 134-149; 185-192, 'X', 194-200, 'X', 202 <SIM>  
 R:Link, P.; Hansen, O.C.; Horn, N.  
 J. Immunol. 140, 4256-4262, 1988  
 A:Title: The binding of mouse hybridoma and human IgE antibodies to the major fecal allergen d by solid-phase inhibition assays with radiolabeled antigen.  
 A:Reference number: A92819; MUID:88229138; PMID:3372999  
 A:Accession: C27634  
 A:Molecule type: protein  
 A:Residues: 24-52 <LIN>  
 C:Superfamily: papain  
 C:Keywords: glycoprotein  
 F:24-245/Product: allergen Der p1 #status predicted <MAT>  
 F:75/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.2%; Score 55; DB 2; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 0.29; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0

Qy 1 RMQGGCGSC 9  
 |||||  
 Db 49 RMQGGCGSC 57  
 |||||

RESULT 3  
 A61500  
 allergen Der f I precursor - house-dust mite (Dermatophagoides farinae)  
 C:Species: Dermatophagoides farinae  
 C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 26-Aug-1999  
 C:Accession: A61500  
 R:DiLworth, R.J.; Chua, K.Y.; Thomas, W.R.  
 Clin. Exp. Allergy 21, 25-32, 1991  
 A:Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I.  
 A:Reference number: A61500; MUID:91215493; PMID:2021874  
 A:Accession: A61500  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-319 <DIL>  
 C:Superfamily: papain

Query Match 90.2%; Score 55; DB 2; Length 319;  
 Best Local Similarity 100.0%; Pred. No. 0.35; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0

Qy 1 RMQGGCGSC 9  
 |||||  
 Db 123 RMQGGCGSC 131  
 |||||

RESULT 4  
 S76753  
 Hypothetical protein - Synchocystis sp. (strain PCC 6803)  
 C:Species: Synchocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S76753  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76753  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-76 <KAN>  
 A:Cross-references: UNIPROT:P74558; EMBL:D90916; GB:AB001339; MID:g1653715; PIDN:BAAL1866  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: conserved hypothetical nifU-like protein HP1492

Query Match 75.4%; Score 46; DB 1; Length 76;  
 Best Local Similarity 77.8%; Pred. No. 2.8;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9  
 |||||  
 Db 39 RLOGACGSC 47  
 |||||

# RESULT 5

AB1970  
 hypothetical protein asrl309 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C:Accession: AB1970

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, R.; Kaneko, T.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AB1970  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-76 <KUR>  
 A:Cross-references: UNIPROT:Q8YX46; GB:BA000019; PIDN:BA073266.1; MID:g17130656; GSPDB:GN  
 A:Experimental source: strain PCC 7120  
 C:Genetics:

C:Superfamily: conserved hypothetical nifU-like protein HP1492

Query Match 75.4%; Score 46; DB 2; Length 76;  
 Best Local Similarity 77.8%; Pred. No. 2.8;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9  
 |||||  
 Db 39 RLOGACGSC 47  
 |||||

# RESULT 6

PQ0650  
 senescence-associated protein SAG2 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
 C:Accession: PQ0650  
 R:Hensel, L.L.; Grbic, V.; Baumgarten, D.A.; Blecker, A.B.  
 Plant Cell 5, 553-564, 1993  
 A:Title: Developmental and age-related processes that influence the longevity and senescence  
 A:Reference number: PQ0650; MUID:93299122; PMID:8518555  
 A:Accession: PQ0650  
 A:Molecule type: mRNA  
 A:Residues: 1-95 <HEN>  
 A:Cross-references: UNIPROT:Q9LL83  
 C:Comment: This protein is a senescence-associated protein.  
 C:Genetics:  
 A:Gene: SAG2  
 C:Superfamily: papain

Qy 1 RMQGGCGSC 9  
 |||||  
 Db 39 RLOGACGSC 47  
 |||||

Query Match 73.8%; Score 45; DB 2; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QGGCGSC 9  
 |||||  
 Db 23 QGGCGSC 29  
 |||||

# RESULT 7

AS9041  
 cysteine proteinase II (EC 3.4.22.-) - ginger



C:Species: Zingiber officinale (ginger)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: A59041  
R:Laursen, R.A.  
A:Submitted to the Protein Sequence Database, July 1999  
A:Description: Amino acid sequences of cysteine proteases from ginger rhizome, Zingiber  
A:Reference number: A59041  
A:Accession: A59041  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-221 <LAU>  
A:Cross-references: UNIPROT:P82474  
A:Experimental source: rhizome  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; glycoprotein; hydrolase  
F:24-65,58-98,155-206/Disulfide bonds: #status experimental  
F:27,161,181/Active site: Cys, His, Asn #status predicted  
F:99,156/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 73.8%; Score 45; DB 2; Length 221;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QGCGGSC 9  
| | | | |  
DB 21 QGCGGSC 27

RESULT 8  
A59040  
Cysteine proteinase I (EC 3.4.22.-) - ginger  
C:Species: Zingiber officinale (ginger)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: A59040  
R:Laursen, R.A.  
A:Submitted to the Protein Sequence Database, July 1999  
A:Description: Amino acid sequences of cysteine proteases from ginger rhizome, Zingiber  
A:Reference number: A59040  
A:Accession: A59040  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-221 <LAU>  
A:Cross-references: UNIPROT:P82473  
A:Experimental source: rhizome  
A:Note: residues 80-87 were not determined but are based on mass measurement and similar  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; glycoprotein; hydrolase  
F:24-65,58-98,155-206/Disulfide bonds: #status predicted  
F:27,161,181/Active site: Cys, His, Asn #status predicted  
F:99,156/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.8%; Score 45; DB 2; Length 221;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QGCGGSC 9  
| | | | |  
DB 21 QGCGGSC 27

RESULT 9  
T08844  
cysteine proteinase (EC 3.4.22.-) isoform B - soybean (fragment)  
N:Alternate names: thiol proteinase isoform B  
C:Species: Glycine max (soybean)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T08844  
R:Herman, E.; Kalinski, A.J.; Rowley, D.  
A:Submitted to the EMBL Data Library, September 1996  
A:Description: Soybean leaf thiol protease isoform B.  
A:Reference number: Z16488  
A:Accession: T08844  
A:Molecule type: mRNA  
A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA  
A:Residues: 1-319 <HER>  
A:Cross-references: UNIPROT:Q96454; EMBL:U71379; NID:g1619902; PID:g1619903  
A:Experimental source: cultivar Century; leaf  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:103-143/Disulfide bonds: #status predicted  
F:112,254,281/Active site: Cys, His, Asn #status predicted  
F:192-248/Disulfide bonds: #status predicted  
F:281,282,283/Binding site: substrate (Asn, Ser, Trp) #status predicted

Query Match 73.8%; Score 45; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QGCGGSC 9  
| | | | |  
DB 106 QGCGGSC 112

RESULT 10  
KRR20A  
oryzain (EC 3.4.22.-) alpha precursor - rice  
C:Species: Oryza sativa (rice)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: JU0388; A40053  
R:Watanabe, H.; Abe, K.; Emori, Y.; Hosoyama, H.; Arai, S.  
A:Submitted to JIPID, May 1991  
A:Reference number: JU0388  
A:Accession: JU0388  
A:Molecule type: DNA  
A:Residues: 1-458 <WAT1>  
A:Cross-references: UNIPROT:P25776  
R:Watanabe, H.; Abe, K.; Emori, Y.; Hosoyama, H.; Arai, S.  
J. Biol. Chem. 266, 16897-16902, 1991  
A:Title: Molecular cloning and gibberellin-induced expression of multiple cysteine protei  
A:Reference number: A40053; MUID:91358494; PMID:1885617  
A:Accession: A40053  
A:Molecule type: mRNA  
A:Residues: 1-458 <WAT2>  
A:Cross-references: GB:D90406; NID:g218180; PID:BAA14402.1; PID:g218181  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; glycoprotein; hydrolase; seed  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-128/Domain: amino-terminal propeptide #status predicted <PRO>  
F:129-348/Product: oryzain alpha #status predicted <MAT>  
F:349-458/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F:150-192,184-225,283-334/Disulfide bonds: #status predicted  
F:153,289,309/Active site: Cys, His, Asn #status predicted  
F:445/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 73.8%; Score 45; DB 1; Length 458;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QGCGGSC 9  
| | | | |  
DB 147 QGCGGSC 153

RESULT 11  
JN0719  
drought-inducible cysteine proteinase (EC 3.4.22.-) RD21A precursor - Arabidopsis thalian  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C:Accession: JN0719  
R:Koizumi, M.; Yamaguchi-Shinozaki, K.; Tsuji, H.; Shinozaki, K.  
Gene 129, 175-182, 1993  
A:Title: Structure and expression of two genes that encode distinct drought-inducible cy  
A:Reference number: JN0718; MUID:93314960; PMID:8325504  
A:Accession: JN0719  
A:Molecule type: mRNA  
A:Residues: 1-462 <KOI>

A;Cross-references: UNIPROT:P43297; GB:D13043; NID:G414340; PIDN:BAA02374.1; PID:G435619  
 C;Comment: This enzyme is involved in the degradation of vegetative storage proteins loc  
 C;Genetics:

A;Gene: rd21A  
 A;Introns: 159/1; 237/3; 284/3; 434/3

C;Superfamily: papain

C;Keywords: cysteine proteinase; hydrolase

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-136/Domain: propeptide #status predicted <PRO>

F;137-462/Product: drought-inducible cysteine proteinase RD21A #status predicted <MAT>

F;155-296/317,318,319/Binding site: substrate (Gln, Asp, Asn, Ser, Trp) #status predicted

F;158-192,233-291/Disulfide bonds: #status predicted

F;161,297,317/Active site: Cys, His, Asn #status predicted

Query Match 73.8%; Score 45; DB 2; Length 462;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QGGCGSC 9

Db 155 QGGCGSC 161

#### RESULT 12

T08153

cysteine proteinase (EC 3.4.22.-) - Volvox carteri f. nagariensis

C;Species: Volvox carteri f. nagariensis

C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 09-Jul-2004

C;Accession: T08153

R;Anon. P.; Haas, E.; Sumper, M.

A;Title: The sex-inducing pheromone and wounding trigger the same set of genes in the mu

A;Reference number: T16381; MUID:98259000; PMID:9596636

A;Accession: T08153

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-658 <AMO>

A;Cross-references: UNIPROT:O65214; EMBL:AF058717; NID:G3047414; PIDN:AAC13728.1; PID:G3

A;Note: sex pheromone and wound-inducible

C;Keywords: cysteine proteinase; hydrolase

Query Match 73.8%; Score 45; DB 2; Length 658;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QGGCGSC 9

Db 232 QGGCGSC 238

#### RESULT 13

T12042

cysteine proteinase (EC 3.4.22.-) 4 precursor - kidney bean

C;Species: Phaseolus vulgaris (kidney bean)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004

C;Accession: T12042

R;Senyuk, V.; Becker, C.; Muentz, K.

A;Title: Isolation of cDNA clone encoding cysteine proteinase (CP4) from a cotyled

A;Reference number: Z17388

A;Accession: T12042

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-377 <SEN>

A;Cross-references: UNIPROT:O24324; EMBL:Z99955

A;Experimental source: cultivar Moldavian; cotyledon; clone cp4

C;Superfamily: papain

C;Keywords: cysteine proteinase; hydrolase

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-137/Domain: propeptide #status predicted <PRO>

F;138-377/Product: cysteine proteinase 4 #status predicted <MAT>

F;162,305,332/Active site: Cys, His, Asn #status predicted

Query Match 72.1%; Score 44; DB 2; Length 377;

Best Local Similarity 77.8%; Pred. No. 17;

Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 RMQGGCGSC 9

Db 154 RMQGGCGSC 162

#### RESULT 14

T04246

hypothetical protein F20B18.20 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T04246

R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.

A;Title: submitted to the Protein Sequence Database, March 1999

A;Reference number: Z15263

A;Accession: T04246

A;Molecule type: DNA

A;Residues: 1-224 <BEV>

A;Cross-references: UNIPROT:Q9SZG6; EMBL:AL049483

A;Experimental source: cultivar Columbia; BAC clone F20B18

C;Genetics:

A;Map position: 4

A;Introns: 22/3; 62/1; 161/3

A;Note: F20B18.20

Query Match 70.5%; Score 43; DB 2; Length 224;

Best Local Similarity 66.7%; Pred. No. 17;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9

Db 106 KMQGGCGSC 116

#### RESULT 15

JC7183

cathepsin Q (EC 3.4.22.-) precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 09-Jul-2004

C;Accession: JC7183

R;Sol-Church, K.; French, J.; Mason, R.W.

A;Title: Cathepsin Q, a novel lysosomal cysteine protease highly expressed in placenta.

A;Reference number: JC7183; MUID:20139729; PMID:10673370

A;Accession: JC7183

A;Molecule type: mRNA

A;Residues: 1-343 <SOL>

A;Cross-references: UNIPROT:Q9QZE3; GB:AF187323; NID:G610770; PIDN:AAF01247.1; PID:G6010

A;Experimental source: placenta

C;Superfamily: papain

C;Keywords: cysteine proteinase; glycoprotein; hydrolase; placenta

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-125/Domain: propeptide #status predicted <PRO>

F;126-343/Product: cathepsin Q #status predicted <MAT>

F;149,286,310/Active site: Cys, His, Asn #status predicted

F;228,299/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.5%; Score 43; DB 2; Length 343;

Best Local Similarity 77.8%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9

Db 141 RMQGGCGSC 149

Search completed: October 27, 2004, 17:51:27

Job time : 4.66932 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:34:27 ; Search time 14.7809 seconds  
(without alignments)

389.270 Million cell updates/sec

Title: US-09-867-159A-3

Perfect score: 61

Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	90.2	210	2 Q9GY0	Q9GY0 dermatophag
2	55	90.2	320	1 MMAL_DERPT	P08176 dermatophag
3	55	90.2	321	1 EUM1_EURMA	P25780 euroglyphus
4	55	90.2	321	1 MMAL_DERFA	P15311 dermatophag
5	55	90.2	321	2 BAC53948	Bac53948 dermatoph
6	49	80.3	337	2 Q6JZ24	Q6JZ24 fundulus he
7	49	80.3	337	2 AA064477	AA064477 fundulus
8	48	78.7	221	2 P95PJ4	Q95PJ4 blomia trop
9	46	75.4	76	2 P74558	P74558 synechocyst
10	46	75.4	76	2 Q8YXA6	Q8YXA6 anabaena sp
11	46	75.4	81	2 Q7U5L6	Q7U5L6 synecococc
12	46	75.4	81	2 Q7V2Q3	Q7V2Q3 prochloroco
13	46	75.4	81	2 Q7V8S7	Q7V8S7 prochloroco
14	46	75.4	81	2 Q7VDF9	Q7VDF9 prochloroco
15	46	75.4	85	2 Q7NCU9	Q7NCU9 gloeobacter
16	46	75.4	170	2 Q81883	Q81883 leptinotars
17	46	75.4	326	2 Q6QRP6	Q6QRP6 leptinotars
18	46	75.4	326	2 Q6QRP7	Q6QRP7 leptinotars
19	46	75.4	326	2 AAS20591	AAS20591 leptinota
20	46	75.4	326	2 AAS20592	AAS20592 leptinota
21	46	75.4	334	2 Q7T183	Q7T183 paralabidoc
22	46	75.4	376	2 Q8GVR2	Q8GVR2 oryza sativ
23	45	73.8	95	2 Q9S9A9	Q9S9A9 arabidopsis
24	45	73.8	181	2 Q84XA1	Q84XA1 carica cand
25	45	73.8	221	1 GPII_ZINOF	P82474 zingiber of
26	45	73.8	221	1 GPII_ZINOF	P82473 zingiber of
27	45	73.8	263	2 Q27059	Q27059 trypanosoma
28	45	73.8	298	2 Q7QR03	Q7QR03 giardia lam
29	45	73.8	299	1 CAL3_GIALA	P92133 giardia lam
30	45	73.8	300	1 CAL2_GIALA	P92132 giardia lam
31	45	73.8	300	2 Q7R189	Q7R189 giardia lam

32 45 73.8 319 2 Q96454 Q96454 glycine max  
33 45 73.8 320 2 Q6DE57 Q6DE57 xenopus lae  
34 45 73.8 326 2 Q24941 Q24941 fasciola he  
35 45 73.8 326 2 Q9NB30 Q9NB30 fasciola he  
36 45 73.8 333 2 Q7T0S4 Q7T0S4 xenopus lae  
37 45 73.8 333 2 Q6DJC1 Q6DJC1 xenopus tro  
38 45 73.8 358 2 Q8H166 Q8H166 arabidopsis  
39 45 73.8 358 2 Q9LL83 Q9LL83 arabidopsis  
40 45 73.8 359 2 Q8W178 Q8W178 brassica ol  
41 45 73.8 365 2 Q84M25 Q84M25 helianthus  
42 45 73.8 370 2 Q9XF79 Q9XF79 sandersonia  
43 45 73.8 421 2 Q9B159 Q9B159 caenorhabdi  
44 45 73.8 458 1 ORYA\_ORISA R25776 oryza sativ  
45 45 73.8 460 2 Q7XY07 Q7XY07 anthurium a

#### ALIGNMENTS

RESULT 1  
Q9GY0 ID Q9GY0 PRELIMINARY; PRT; 210 AA.  
AC Q9GY0;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Allergen Der f1 (Fragment).  
OS Dermatophagoides farinae (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;  
OC Pyroglyphidae; Dermatophagoides.  
OX NCBI\_TaxID=6954;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hao M.Q., Xu J., Zhong N.S.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; AF285763; AAGO520.1; -;  
DR PIR; A27634; A27634.  
DR GO; GO:0004197; P: cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PR00705; PAPAIN\_C1; 1.  
DR SMART; SMC0645; Pept C1; 1.  
DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; UNKNOWN\_1.  
KW Hydrolase; Protease; Thiol protease.  
FT NON TER 1  
FT NON TER 210  
SQ SEQUENCE 210 AA; 23548 MW; BA08029D642EEB90 CRC64;  
Query Match 90.2%; Score 55; DB 2; Length 210;  
Best Local Similarity 100.0%; Pred. No. 0.82;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RMQGGCGSC 9  
Db 26 RMQGGCGSC 34  
RESULT 2  
MMAL\_DERPT ID MMAL\_DERPT STANDARD; PRT; 320 AA.  
AC P08176; Q24616;  
DT 01-AUG-1998 (Rel. 08, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Major mite fecal allergen Der p 1 precursor (EC 3.4.22.-) (Der p I).  
GN Name=DERP1;  
OS Dermatophagoides pteronyssinus (House-dust mite).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcophagidae; Ascarididae; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Dermatophagoides.  
 RX NCBI\_TaxID=6956;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND POLYMORPHISM.  
 RX MEDLINE=93357682; PubMed=8353459;  
 RA Chua K.Y., Kehal P.K., Thomas W.R.;  
 RT "Sequence polymorphisms of cDNA clones encoding the mite allergen Der  
 p 1.";  
 RL Int. Arch. Allergy Immunol. 101:364-368(1993).  
 RN [2]  
 RP SEQUENCE OF 76-320 FROM N.A.  
 RX MEDLINE=88089411; PubMed=3335830;  
 RA Chua K.Y., Stewart G.A., Thomas W.R., Simpson R.J., Dilworth R.J.,  
 Plozza T.M., Turner K.J.;  
 RT "Sequence analysis of cDNA coding for a major house dust mite  
 allergen, Der p 1. Homology with cysteine proteases.";  
 RL J. Exp. Med. 167:175-182(1988).  
 RN [3]  
 RP SEQUENCE OF 81-176 FROM N.A.  
 RX MEDLINE=8814080; PubMed=3276629;  
 RA Thomas W.R., Stewart G.A., Simpson R.J., Chua K.Y., Plozza T.M.,  
 Dilworth R.J., Nisbet A., Turner K.J.;  
 RT "Cloning and expression of DNA coding for the major house dust mite  
 allergen Der p 1 in Escherichia coli.";  
 RL Int. Arch. Allergy Appl. Immunol. 95:127-129(1988).  
 RN [4]  
 RP REVISIONS TO 232-241.  
 RX MEDLINE=91215493; PubMed=2021874;  
 RA Dilworth R.J., Chua K.Y., Thomas W.R.;  
 RT "Sequence analysis of cDNA coding for a major house dust mite  
 allergen, Der p 1.";  
 RL Clin. Exp. Allergy 21:25-32(1991).  
 RN [5]  
 RP SEQUENCE OF 99-308 FROM N.A.  
 RX MEDLINE=93130112; PubMed=1493062;  
 RA Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
 RT "Molecular characterisation of group I allergen Eur m I from house  
 dust mite Euroglyphus maynei.";  
 RL Int. Arch. Allergy Immunol. 99:150-152(1992).  
 RN [6]  
 RP SEQUENCE OF 99-127.  
 RX MEDLINE=88229138; PubMed=3372999;  
 RA Lind P., Hansen O.C., Horn N.;  
 RT "The binding of mouse hybridoma and human IgE antibodies to the major  
 fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative  
 binding site location and species specificity studied by solid-phase  
 inhibition assays with radio-labeled antigen.";  
 RL J. Immunol. 140:4256-4262(1988).  
 RN [7]  
 RP SEQUENCE OF 99-139; 177-192; 208-224 AND 260-277, AND VARIANT ALA-222.  
 RX PubMed=2911558;  
 RA Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;  
 RT "Structural studies on the allergen Der p 1 from the house dust mite  
 Dermatophagoides pteronyssinus: similarity with cysteine  
 proteinases.";  
 RL Protein Seq. Data Anal. 2:17-21(1989).  
 RN [8]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE=95062135; PubMed=7971950;  
 RA Topham C.M., Srinivasan N., Thorpe C.J., Overington J.P.,  
 Kalsheker N.A.;  
 RT "Comparative modelling of major house dust mite allergen Der p 1:  
 structure validation using an extended environmental amino acid  
 propensity table.";  
 RL Protein Eng. 7:869-894(1994).  
 CC -1- FUNCTION: Thiol protease that hydrolyzes proteins, with a  
 preference for phe or basic residues.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
 mite allergy are bronchial asthma, allergic rhinitis and  
 conjunctivitis. Reacts with IgE in 80% of patients with house dust

CC allergy.  
 CC -1- SIMILARITY: Belongs to peptidase family C1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; U11695; AAB60215.1;  
 DR EMBL; M24794; AAA28296.1; ALT\_INIT.  
 DR EMBL; X65197; CAA46317.1;  
 DR PIR; JQ0337; JQ0337.  
 DR HSP; P53634; IK3B.  
 DR MEROPS; C01.073;  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; Pept\_cys\_acsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPA1N; 1.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SMO0645; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
 DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 KW Allergen; Direct protein sequencing; Glycoprotein; Hydrolase;  
 KW Polymorphism; Signal; Thiol protease; Zymogen.  
 FT SIGNAL 1 18 Potential.  
 FT PROPEP 19 98 Activation peptide.  
 FT CHAIN 99 320 Major mite fecal allergen Der p 1.  
 FT CARBOHYD 150 150 N-linked (GlcNAc...) (Potential).  
 FT ACT\_SITE 132 132 By similarity.  
 FT ACT\_SITE 268 268 By similarity.  
 FT ACT\_SITE 288 288 By similarity.  
 FT DISULFID 102 215 By similarity.  
 FT DISULFID 129 169 By similarity.  
 FT DISULFID 163 201 By similarity.  
 FT VARIANT 148 148 Y -> H.  
 FT VARIANT 179 179 E -> K.  
 FT VARIANT 222 222 V -> A.  
 FT VARIANT 234 234 S -> T.  
 FT VARIANT 313 313 E -> Q.  
 SQ SEQUENCE 320 AA; 36104 MW; A0B1F4DD09791DFE CRC64;  
 Query Match 90.2%; Score 55; DB 1; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RMGGGCGSC 9  
 Db 124 RMGGGCGSC 132  
 RESULT 3  
 EMBL EMBL  
 ID EMBL EMBL STANDARD; PRT; 321 AA.  
 AC P25780; Q9TZZ3; Q9TZZ4; Q9UBA0;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Mite group 1 allergen Eur m 1 precursor (EC 3.4.22.-) (Eur m I).  
 GN Name=EURM1;  
 OS Euroglyphus maynei (Wayne's house dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcophagidae; Ascarididae; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Dermatophagoides.  
 OC NCBI\_TaxID=6958;  
 RN [1]  
 RP SEQUENCE FROM N.A. (EUR M 1.0101 AND EUR M 1.0102).  
 RX MEDLINE=93126275; PubMed=925958;  
 RA Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;  
 RT "Molecular analysis of the group 1 and 2 allergens from the house dust

mte, Euroglyphus maynei.;  
Int. Arch. Allergy Immunol. 118:15-22(1999).  
[2]  
SEQUENCE OF 99-309 FROM N.A.  
MEDLINE=93130112; PubMed=1483062;  
Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
"molecular characterisation of group I allergen Eur m I from house  
dust mite Euroglyphus maynei.";  
Int. Arch. Allergy Immunol. 99:150-152(1992).  
Int. Arch. Allergy Immunol. 99:150-152(1992).  
-!- FUNCTION: Probable thiol protease.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
mite allergy are bronchial asthma, allergic rhinitis and  
conjunctivitis.  
-!- SIMILARITY: Belongs to peptidase family C1.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; AF047610; AAC82351.1; -;  
EMBL; AF047611; AAC82352.1; ALT INIT.  
EMBL; AF047612; AAC82353.1; -;  
EMBL; X60073; CAA42677.1; -;  
PIR; S21864; S21864.  
HSSP; P53634; 1K3B.  
MEROPS; C01.073; -;  
InterPro; IPR000688; Peptidase\_C1.  
InterPro; IPR000169; Pept\_cys\_acsite.  
Pfam; PF00112; Peptidase\_C1; 1.  
PRINTS; PR00705; PAPAIN.  
ProDom; PD000158; Peptidase\_C1; 1.  
SMART; SM00645; Pept\_C1; 1.  
PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
Allergen; Glycoprotein; Hydrolase; Signal; Thiol protease; Zymogen.  
FT SIGNAL 1 18 Potential.  
FT PROPEP 19 98  
FT CHAIN 99 321 Mite group 1 allergen Eur m 1.  
FT ACT\_SITE 133 133 By similarity.  
FT ACT\_SITE 269 289 By similarity.  
FT ACT\_SITE 289 289 By similarity.  
FT DISULFID 130 170 By similarity.  
FT CARBOHYD 34 34 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).  
FT VARIANT 36 36 T -> S (in Eur m 1.0102).  
FT VARIANT 126 126 M -> N (in Eur m 1.0102).  
FT VARIANT 320 320 M -> I (in Eur m 1.0102).  
SQ SEQUENCE 321 AA; 36290 MW; 6CFD44FEC725999E CRC64;  
-----  
Query Match 90.2%; Score 55; DB 1; Length 321;  
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
QY 1 RMGGGCGSC 9  
Db 125 RMGGGCGSC 133  
-----  
RESULT 4  
MMAL\_DERFA STANDARD; PRT; 321 AA.  
AC P16311;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Major mite fecal allergen Der f 1 precursor (EC 3.4.22.-) (Der f 1).  
GN Name=DERf1;

OS  
CC Dermaphagoides farinae (House-dust mite).  
CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
CC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;  
CC Pyroglyphidae; Dermatophagoides.  
OX NCBI\_TaxID=6954;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=91215493; PubMed=2021874;  
RA Diworth R.J., Chua K.Y., Thomas W.R.;  
RT "Sequence analysis of cDNA coding for a major house dust mite  
allergen, Der f 1.";  
RL Clin. Exp. Allergy 21:25-32(1991).  
RN [2]  
RN SEQUENCE OF 98-309 FROM N.A.  
RA Kent N., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 99-128.  
RX MEDLINE=88229138; PubMed=3372999;  
RA Lind P., Hansen O.C., Horn N.;  
RT "The binding of mouse hybridoma and human IgE antibodies to the major  
fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative  
binding site location and species specificity studied by solid-phase  
inhibition assays with radiolabeled antigen.";  
RL J. Immunol. 140:4256-4262(1988).  
CC -!- FUNCTION: Thiol protease that hydrolyzes proteins, with a  
preference for Phe or basic residues.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
mite allergy are bronchial asthma, allergic rhinitis and  
conjunctivitis.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL; X65196; CAA46316.1; -;  
PIR; A27634; A27634.  
HSSP; P53634; 1K3B.  
MEROPS; C01.073; -;  
InterPro; IPR000688; Peptidase\_C1.  
InterPro; IPR000169; Pept\_cys\_acsite.  
Pfam; PF00112; Peptidase\_C1; 1.  
PRINTS; PR00705; PAPAIN.  
ProDom; PD000158; Peptidase\_C1; 1.  
SMART; SM00645; Pept\_C1; 1.  
PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
Allergen; Direct protein sequencing; Glycoprotein; Hydrolase; Signal;  
Thiol protease; Zymogen.  
FT SIGNAL 1 18 Potential.  
FT PROPEP 19 98 Activation peptide.  
FT CHAIN 99 321 Major mite fecal allergen Der f 1.  
FT ACT\_SITE 133 133 By similarity.  
FT ACT\_SITE 269 289 By similarity.  
FT ACT\_SITE 288 288 By similarity.  
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).  
FT DISULFID 102 216 By similarity.  
FT DISULFID 130 170 By similarity.  
FT DISULFID 164 202 By similarity.  
FT CONFLICT 201 201 R -> Q (in Ref. 2).  
FT CONFLICT 282 282 D -> V (in Ref. 2).  
SQ SEQUENCE 321 AA; 36435 MW; 04523E54EBB476E CRC64;  
-----  
Query Match 90.2%; Score 55; DB 1; Length 321;  
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 RMGGCGSC 9
Db 125 RMGGCGSC 133

RESULT 5
ID BAC53948 PRELIMINARY; PRT; 321 AA.
AC BAC53948
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Der f 1 allergen preproenzyme precursor.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6954;
[1]
RN SEQUENCE FROM N.A.
RP Yasuhara T., Takai T., Takahashi K., Yuuki T., Yokota T., Okudaira H.,
RA Okumura Y.;
RT "Dermatophagoides farinae Der f 1 allergen preproenzyme mRNA.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Dilworth R.J., Chuan K.Y., Thomas W.R.;
RT "Sequence analysis of cDNA coding for a major house dust mite
allergen, Der f 1";
RL Clin. Exp. Allergy 21:25-32(1991).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=93357682; PubMed=8353459;
RA Chua K.Y., Kehal P.K., Thomas W.R.;
RT "Sequence polymorphisms of cDNA clones encoding the mite allergen Der
p 1";
RL Int. Arch. Allergy Immunol. 101:364-368(1993).
DR EMBL; AB034946; BAC53948.1; -
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 99 321
FT VARIANT 201 201 Q -> R (IN REF. 2).
FT VARIANT 282 282 V -> D (IN REF. 2).
SQ SEQUENCE 321 AA; 36391 MW; 83594754EBB4477 CRC64;

Query Match 90.2%; Score 55; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMGGCGSC 9
Db 125 RMGGCGSC 133

RESULT 6
Q6JZZ4 PRELIMINARY; PRT; 337 AA.
AC Q6JZZ4
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Cathepsin S.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
[1]
RN SEQUENCE FROM N.A.
RP PubMed=14735489;
RA Fabra M., Cerdà J.;
RT "Ovarian cysteine proteinases in the teleost Fundulus heteroclitus:
molecular cloning and gene expression during vitellogenesis and oocyte
maturation.";
RL Mol. Reprod. Dev. 67:282-294(2004).
DR EMBL; AY225486; AAC64477.1; -
SQ SEQUENCE 337 AA; 37390 MW; BE3348A5B93DBE6B CRC64;

Query Match 80.3%; Score 49; DB 2; Length 337;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RMGGCGSC 9
Db 138 RMGGCGSC 146

RESULT 7
AAC64477 PRELIMINARY; PRT; 337 AA.
ID AAC64477
AC AAC64477
DT 24-MAY-2004 (TReMBLrel. 27, Created)
DT 24-MAY-2004 (TReMBLrel. 27, Last sequence update)
DT 24-MAY-2004 (TReMBLrel. 27, Last annotation update)
DE Cathepsin S.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
[1]
RN SEQUENCE FROM N.A.
RP PubMed=14735489;
RA Fabra M., Cerdà J.;
RT "Ovarian cysteine proteinases in the teleost Fundulus heteroclitus:
molecular cloning and gene expression during vitellogenesis and oocyte
maturation.";
RL Mol. Reprod. Dev. 67:282-294(2004).
DR EMBL; AY225486; AAC64477.1; -
SQ SEQUENCE 337 AA; 37390 MW; BE3348A5B93DBE6B CRC64;

Query Match 80.3%; Score 49; DB 2; Length 337;
Best Local Similarity 88.9%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RMGGCGSC 9
Db 138 RMGGCGSC 146

RESULT 8
Q95PJ4 PRELIMINARY; PRT; 221 AA.
ID Q95PJ4
AC Q95PJ4
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Cysteine protease (Fragment).
OS Blomia tropicalis (Mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Glycyphagoidea;
OC Echimyopodidae; Blomia.

```

```

OX NCBI_TaxID=40697;
RN [1]
RP SEQUENCE FROM N.A.
RA Mora C.I., Diaz A.M., Montealegre F., Flores I.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF277840; AAK58415.1; -.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPA1N.
DR SMART; SM00645; Pept_C1; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Protease.
FT NON_TER
FT CHAIN
FT SEQUENCE 221 AA; 25126 MW; 272B45ESA53F2900 CRC64;
SQ
Query Match 78.7%; Score 48; DB 2; Length 221;
Best Local Similarity 88.9%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RMOGGCGSC 9
DB 17 RMOGGCGSC 25
RESULT 9
P74558 PRELIMINARY; PRT; 76 AA.
AC P74558;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nifu protein.
GN OrderedLocusNames=ss12667;
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. I. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugiyura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
DR EMBL; D90916; BAAL8665.1; -.
DR FIR; S76753; S76753.
DR InterPro; IPR001075; Nifu_C.
DR Pfam; PF01106; Nifu; 1.
DR ProDom; PD002830; Nifu_C; 1.
KW Complete proteome.
SQ SEQUENCE 76 AA; 8374 MW; C58AE2B89F856D72 CRC64;
Query Match 75.4%; Score 46; DB 2; Length 76;
Best Local Similarity 77.8%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 RMOGGCGSC 9
DB 39 RMOGGCGSC 47
RESULT 11
Q7U5L6 PRELIMINARY; PRT; 81 AA.
ID Q7U5L6
AC Q7U5L6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nifu-like protein.
GN OrderedLocusNames=SYNW1687;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarrren J.,
RA Paulsen I.F., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569693; CAE08202.1; -.
DR InterPro; IPR001075; Nifu_C.
DR Pfam; PF01106; Nifu; 1.
DR ProDom; PD002830; Nifu_C; 1.
KW Complete proteome.
SQ SEQUENCE 81 AA; 8837 MW; DCAA0CD4F48DC7CD CRC64;

```

```

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 RMOGGCGSC 9
DB 39 RMOGGCGSC 47
RESULT 10
Q8YXA6 PRELIMINARY; PRT; 76 AA.
ID Q8YXA6
AC Q8YXA6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Asr1309 protein.
GN OrderedLocusNames=asr1309;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpou S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003595; BAB73266.1; -.
DR FIRM; AB1970; AB1970.
DR InterPro; IPR001075; Nifu_C.
DR Pfam; PF01106; Nifu; 1.
DR ProDom; PD002830; Nifu_C; 1.
KW Complete proteome.
SQ SEQUENCE 76 AA; 8475 MW; D7E23575AA96C1D4 CRC64;
Query Match 75.4%; Score 46; DB 2; Length 76;
Best Local Similarity 77.8%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 RMOGGCGSC 9
DB 39 RMOGGCGSC 47
RESULT 11
Q7U5L6 PRELIMINARY; PRT; 81 AA.
ID Q7U5L6
AC Q7U5L6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nifu-like protein.
GN OrderedLocusNames=SYNW1687;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarrren J.,
RA Paulsen I.F., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569693; CAE08202.1; -.
DR InterPro; IPR001075; Nifu_C.
DR Pfam; PF01106; Nifu; 1.
DR ProDom; PD002830; Nifu_C; 1.
KW Complete proteome.
SQ SEQUENCE 81 AA; 8837 MW; DCAA0CD4F48DC7CD CRC64;

```

Query Match 75.4%; Score 46; DB 2; Length 81;  
Best Local Similarity 77.8%; Pred. No. 8.8;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9  
Db 44 RLQAGCGSC 52

## RESULT 12

ID Q7V2Q3 PRELIMINARY; PRT; 81 AA.  
AC Q7V2Q3;  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE Nifu-like protein.  
GN OrderedLocusNames=PM00418;  
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=59919;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;  
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,  
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,  
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
RA Webb E.A., Zinser E.R., Chisholm S.W.;  
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation."  
RL Nature 424:1042-1047(2003).  
DR EMBL: BX572091; CAE18877.1;  
DR InterPro: IPR001075; Nifu\_C.  
DR Pfam: PF01106; Nifu; 1.  
DR ProDom: PD002830; Nifu\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 81 AA; 8815 MW; 3D59B4218DD46C05 CRC64;

Query Match 75.4%; Score 46; DB 2; Length 81;  
Best Local Similarity 77.8%; Pred. No. 8.8;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9  
Db 44 RLQAGCGSC 52

## RESULT 13

ID Q7V8S7 PRELIMINARY; PRT; 81 AA.  
AC Q7V8S7;  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE Nifu-like protein.  
GN OrderedLocusNames=PM00254;  
OS Prochlorococcus marinus (strain MIT 9313).  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=74547;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;  
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,  
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,  
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,  
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,  
RA Webb E.A., Zinser E.R., Chisholm S.W.;  
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation."

RL Nature 424:1042-1047(2003).  
DR EMBL: BX572095; CAE20429.1;  
DR InterPro: IPR001075; Nifu\_C.  
DR Pfam: PF01106; Nifu; 1.  
DR ProDom: PD002830; Nifu\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 81 AA; 8807 MW; E433B9C45B8A515A CRC64;

Query Match 75.4%; Score 46; DB 2; Length 81;  
Best Local Similarity 77.8%; Pred. No. 8.8;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9  
Db 44 RLQAGCGSC 52

## RESULT 14

ID Q7VDF9 PRELIMINARY; PRT; 81 AA.  
AC Q7VDF9;  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE Thioredoxin family protein.  
GN OrderedLocusNames=Pro0417;  
OS Prochlorococcus marinus.  
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
OC Prochlorococcus.  
OX NCBI\_TaxID=1219;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SARG / CCMP 1375 / SS120;  
RX MEDLINE=22810154; PubMed=12917486;  
RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,  
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,  
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,  
RA Scanlan D.J., Tandeau de Marsac N., Weissbach J., Wincker P.,  
RA Wolf Y.I., Hess W.R.;  
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120, a nearly minimal oxypototrophic genome."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).  
DR EMBL: AE017162; AAP9463.1;  
DR InterPro: IPR001075; Nifu\_C.  
DR Pfam: PF01106; Nifu; 1.  
DR ProDom: PD002830; Nifu\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 81 AA; 8858 MW; 86FC9C746AAD88A0 CRC64;

Query Match 75.4%; Score 46; DB 2; Length 81;  
Best Local Similarity 77.8%; Pred. No. 8.8;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9  
Db 44 RLQAGCGSC 52

## RESULT 15

ID Q7NCU9 PRELIMINARY; PRT; 85 AA.  
AC Q7NCU9;  
DT 01-MAR-2004 (TREMELrel. 26, Created)  
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)  
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)  
DE Gsl2877 protein.  
GN OrderedLocusNames=gsl2877;  
OS Gloeobacter violaceus.  
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.  
OX NCBI\_TaxID=33072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=PCC 7421;



RX MEDLINE=22977040; PubMed=14621292;  
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,  
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,  
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a  
RT cyanobacterium that lacks thylakoids.";  
RL DNA Res. 10:137-145(2003)  
DR EMBL; AF065578; BAC90818.1; -.  
DR InterPro; IPR001075; NifU\_C.  
DR Pfam; PF01106; NifU; 1.  
DR ProDom; PD002830; NifU\_C; 1.  
KW Complete proteome.  
SQ SEQUENCE 85 AA; 9590 MW; BF28726CE6FCAB00 CRC64;

Query Match 75.4%; Score 46; DB 2; Length 85;  
Best Local Similarity 77.8%; Pred. No. 9.2;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMQGGGSC 9  
|:|:|:|:|  
Db 48 RLQAGGSC 56

Search completed: October 27, 2004, 17:50:16  
Job time : 16.7809 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:26:32 ; Search time 13.7849 Seconds  
(without alignments)  
260.234 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMQGGGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%  
Maximum Watch 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04: \*  
1: Geneseq1980s: \*  
2: Geneseq1990s: \*  
3: Geneseq2000s: \*  
4: Geneseq2001s: \*  
5: Geneseq2002s: \*  
6: Geneseq2003as: \*  
7: Geneseq2003bs: \*  
8: Geneseq2004s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	100.0	10	5 AAO20569	Aao20569 Cysteine
2	61	100.0	10	5 ABB98534	Abb98534 Cysteine
3	55	90.2	27	2 AAR34687	Aar34687 DFI-2 (13-
4	55	90.2	27	2 AAR36447	Aar36447 DFI-2.1(1
5	55	90.2	27	2 AAR51732	Aar51732 Der p I d
6	55	90.2	27	2 AAR51735	Aar51735 Der f I d
7	55	90.2	27	2 AAR77161	Aar77161 Dermatoph
8	55	90.2	27	2 AAW71913	Aaw71913 Dermatoph
9	55	90.2	27	2 AAW71970	Aaw71970 Dermatoph
10	55	90.2	27	2 AAY50361	Aay50361 Dermatoph
11	55	90.2	27	2 AAY50424	Aay50424 Dermatoph
12	55	90.2	27	4 AAU19027	Aau19027 T-cell ep
13	55	90.2	27	4 AAU18964	Aau18964 T-cell ep
14	55	90.2	29	2 AAR36448	Aar36448 DFI-3 (21-
15	55	90.2	29	2 AAR34688	Aar34688 Der f I d
16	55	90.2	29	2 AAR51736	Aar51736 Der p I d
17	55	90.2	29	2 AAR51733	Aar51733 Dermatoph
18	55	90.2	29	2 AAW71913	Aaw71913 Dermatoph
19	55	90.2	29	2 AAW71971	Aaw71971 Dermatoph
20	55	90.2	29	2 AAW71914	Aaw71914 Dermatoph
21	55	90.2	29	2 AAY50425	Aay50425 Dermatoph
22	55	90.2	29	2 AAY50362	Aay50362 Dermatoph
23	55	90.2	29	4 AAU19028	Aau19028 T-cell ep
24	55	90.2	29	4 AAU18965	Aau18965 T-cell ep
25	55	90.2	211	2 AAY25677	Aay25677 Euroglyph

26	55	90.2	211	2 AAY25676	Aay25676 Euroglyph
27	55	90.2	211	2 AAY25678	Aay25678 Euroglyph
28	55	90.2	211	7 ADC34925	Adc34925 Euroglyph
29	55	90.2	211	7 ADC34926	Adc34926 Euroglyph
30	55	90.2	211	7 ADC34924	Adc34924 Euroglyph
31	55	90.2	212	2 AAY25679	Aay25679 Euroglyph
32	55	90.2	212	7 ADC34927	Adc34927 Euroglyph
33	55	90.2	222	2 AAR52742	Aar52742 Protein a
34	55	90.2	222	4 AAU07746	Aau07746 House dus
35	55	90.2	222	4 AAU07748	Aau07748 House dus
36	55	90.2	222	4 AAU07747	Aau07747 House dus
37	55	90.2	222	5 AAB98347	Aab98347 D. pteron
38	55	90.2	222	5 ABG67025	Abg67025 House dus
39	55	90.2	222	5 ABG67032	Abg67032 House dus
40	55	90.2	222	5 ABG67027	Abg67027 House dus
41	55	90.2	222	5 ABG67033	Abg67033 House dus
42	55	90.2	222	5 ABG67026	Abg67026 House dus
43	55	90.2	222	5 ABG67029	Abg67029 House dus
44	55	90.2	222	5 ABG67028	Abg67028 House dus
45	55	90.2	222	5 ABG67024	Abg67024 House dus

ALIGNMENTS

RESULT 1  
AAO20569  
ID AAO20569 standard; peptide; 10 AA.  
XX

AC AAC20569;  
XX

DT 02-JAN-2003 (first entry)  
XX

DE Cysteine protease epitope peptide region, SEQ ID No 3.  
XX

KW Antiallergic; antiinflammatory; antiashmatic; dermatologic; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopic eczema; epitope.  
XX

OS Dermatophagoides pteronyssinus.  
XX

PN WO200278736-A2.  
XX

PD 10-OCT-2002.  
XX

PF 28-MAR-2002; 2002WO-FR001098.  
XX

PR 30-MAR-2001; 2001FR-00004370.  
XX

PR 03-MAY-2001; 2001FR-00005929.  
XX

PR 29-MAY-2001; 2001US-00867159.  
XX

PA (ANTI-) ANTIALIS SARL.  
XX

PI Loria E, Terrasse G, Trehin Y;  
XX

WPI; 2002-750636/81.  
XX

XX Antiallergic compositions containing an anti-histamine, a histamine  
PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
for the allergen.

PS Claim 14; Page 11; 32pp; French.  
XX

CC The invention relates to antiallergic compositions containing an anti-  
histamine, a histamine synthesis inhibitor, and optionally an allergen or  
isolated nucleic acid molecule that has at least one polynucleotide  
sequence coding for the allergen, together with a pharmaceutical carrier.  
CC The pharmaceutical composition of the invention is useful as a non-  
specific antiallergic treatment, and also useful in the treatment of  
allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
allergic and atopic eczema. This sequence represents a peptide of a  
cysteine protease epitope region relating to the antiallergic

```

CC compositions of the invention
XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 61; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSCN 10
Db 1 RMQGGCGSCN 10

RESULT 2
ABB98534
ID ABB98534 standard; peptide; 10 AA.
XX
AC ABB98534;
XX
DT 13-DEC-2002 (first entry)
XX
DE Cysteine protease epitope #1.
XX
KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunotherapy; allergen; allergic hypersensitivity reaction;
KW allergic asthma; allergic rhinitis; allergic atopic eczema;
KW cysteine protease.
XX
OS Dermatophagoides pteronyssinus.
XX
FN FR2822709-A1.
XX
PD 04-OCT-2002.
XX
PF 03-MAY-2001; 2001FR-00005929.
XX
PR 30-MAR-2001; 2001FR-00004370.
XX
PA (ANTI-) ANTIALIS SARL.
XX
PI Loria E, Terrasse G, Trehin Y;
XX
DR WPI; 2002-735037/80.
XX
PT Antiallergic composition, useful for preventing and treating e.g. asthma,
PT rhinitis or eczema, containing at least two of allergen, antihistamine
PT and histamine synthesis inhibitor.
XX
PS Claim 8; Page 6; 33pp; French.
XX
CC The present invention relates to an antiallergic pharmaceutical
CC composition (I) comprising a pharmaceutical carrier containing an active
CC agent combination of at least two of: an allergen; an antihistamine; and
CC a histamine synthesis inhibitor. (I) is used for treating or preventing
CC allergic hypersensitivity reactions, especially allergic asthma, allergic
CC rhinitis or allergic atopic eczema, in babies, children or adults. The
CC present sequence is a peptide fragment (epitope) of cysteine protease
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the
CC invention
XX
SQ Sequence 10 AA;
Query Match 100.0%; Score 61; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSCN 10
Db 1 RMQGGCGSCN 10

RESULT 3
AAR34687
ID AAR34687 standard; peptide; 27 AA.
XX
AC AAR34687;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
DE DFI-2.1(13-39), a Dermatophagoides protein allergen.
XX
KW T cell epitope; house dust mite; allergy; soluble; Def pi.
XX
OS Synthetic.
XX
FN WO9308279-A1.
XX
PD 29-APR-1993.
XX
PF 15-OCT-1992; 92WO-US008637.
XX
PR 16-OCT-1991; 91US-00777859.
PR 08-MAY-1992; 92US-00861396.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
DR WPI; 1993-152472/18.
XX
PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis
PT and treatment of sensitivity to house dust mite.
XX
PS Claim 10; Fig 3; 176pp; English.
XX
CC The peptide is one of a series of overlapping peptides synthesised by
CC standard techniques to cover the whole Dermatophagoides pteronyssinus Der
CC pi sequence. The T cell epitopes of the protein were mapped by detection
CC of the peptide's ability to stimulate T cell activity. The peptides may
CC be used for diagnosis and treatment of sensitivity to house dust mite
CC allergens. When administered to house dust mite sensitive individuals,
CC the peptides are capable of modifying the allergic response to the
CC allergens. The peptides may be modified for e.g. increasing solubility,
CC enhancing therapeutic or preventive efficacy or stability. See also
CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct FN
CC field.)
XX
SQ Sequence 27 AA;
Query Match 90.2%; Score 55; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9
Db 14 RMQGGCGSC 22

RESULT 4
AAR36447
ID AAR36447 standard; peptide; 27 AA.
XX
AC AAR36447;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
DE DFI-2.1(13-39), a Dermatophagoides protein allergen.
XX
KW T cell epitope; house dust mite; allergy; soluble; Def pi.
XX
OS Synthetic.
XX
FN WO9308279-A1.
XX

```

29-APR-1993.

15-OCT-1992; 92WO-US008637.

16-OCT-1991; 91US-00777859.

08-MAY-1992; 92US-00881396.

(IMMU-) IMMULOGIC PHARM CORP.

Garman RD, Greenstein JL, Kuo MC, Rogers BL; WPI; 1993-152472/18.

Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis and treatment of sensitivity to house dust mite.

Claim 4; Fig 4; 176pp; English.

The peptide is one of a series of overlapping peptides synthesized by standard techniques to cover the whole Dermatophagoides farinae Def p1 sequence. The T cell epitopes of the protein were mapped by detection of the peptide's ability to stimulate T cell activity. The peptides may be used for diagnosis and treatment of sensitivity to house dust mite allergens. When administered to house dust mite sensitive individuals, the peptides are capable of modifying the allergic response to the allergens. The peptides may be modified for e.g. increasing solubility, enhancing therapeutic or preventive efficacy or stability. See also AAR34686-700 and AAR36399-490. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 27 AA;

Query Match 90.2%; Score 55; DB 2; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.44;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9  
| | | | |  
DB 14 RMQGGCGSC 22

RESULT 5  
AAR51732  
ID AAR51732 standard; protein; 27 AA.  
AC AAR51732;  
DT 01-FEB-1995 (first entry)  
DE Der p I derived peptide, DP I-2(13-39).

Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I; homolog; D. farinae; Der f I; group II; Der f II; T-cell; epitopes; fusion peptides; antigenic fragments; substitution; deletion; addition; chemical synthesis; chemical cleavage; recombinant techniques; allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis; IgE-mediated responses; anergise; lymphokine secretion profile; modify; T cell subpopulations; unresponsive; immune response; tolerise.  
Dermatophagoides pteronyssinus.  
ZA9302677-A.  
26-JAN-1994.  
16-APR-1993; 93ZA-00002677.  
16-APR-1993; 93ZA-00002677.  
(IMMU-) IMMULOGIC PHARM CORP.  
Garman RD, Greenstein JL, Kuo M, Rogers BL;

29-APR-1993.

15-OCT-1992; 92WO-US008637.

16-OCT-1991; 91US-00777859.

08-MAY-1992; 92US-00881396.

(IMMU-) IMMULOGIC PHARM CORP.

Garman RD, Greenstein JL, Kuo MC, Rogers BL; WPI; 1993-152472/18.

Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis and treatment of sensitivity to house dust mite.

Claim 4; Fig 4; 176pp; English.

The peptide is one of a series of overlapping peptides synthesized by standard techniques to cover the whole Dermatophagoides farinae Def p1 sequence. The T cell epitopes of the protein were mapped by detection of the peptide's ability to stimulate T cell activity. The peptides may be used for diagnosis and treatment of sensitivity to house dust mite allergens. When administered to house dust mite sensitive individuals, the peptides are capable of modifying the allergic response to the allergens. The peptides may be modified for e.g. increasing solubility, enhancing therapeutic or preventive efficacy or stability. See also AAR34686-700 and AAR36399-490. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 27 AA;

Query Match 90.2%; Score 55; DB 2; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.44;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9  
| | | | |  
DB 14 RMQGGCGSC 22

RESULT 5  
AAR51732  
ID AAR51732 standard; protein; 27 AA.  
AC AAR51732;  
DT 01-FEB-1995 (first entry)  
DE Der p I derived peptide, DP I-2(13-39).

Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I; homolog; D. farinae; Der f I; group II; Der f II; T-cell; epitopes; fusion peptides; antigenic fragments; substitution; deletion; addition; chemical synthesis; chemical cleavage; recombinant techniques; allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis; IgE-mediated responses; anergise; lymphokine secretion profile; modify; T cell subpopulations; unresponsive; immune response; tolerise.  
Dermatophagoides pteronyssinus.  
ZA9302677-A.  
26-JAN-1994.  
16-APR-1993; 93ZA-00002677.  
16-APR-1993; 93ZA-00002677.  
(IMMU-) IMMULOGIC PHARM CORP.  
Garman RD, Greenstein JL, Kuo M, Rogers BL;

PI Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX WPI; 1994-126807/15.  
 XX Isolated and/or modified peptides comprising T-cell epitopes - of major  
 PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
 PT sensitivity to house dust mites.  
 XX Disclosure; Fig 4; 154pp; English.  
 XX  
 CC The sequences given in AAR51731-841 represent T-cell epitopes derived  
 CC from the group I and II protein allergens from the house dust mite D.  
 CC farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II  
 CC respectively. The Der f II proteinsCC shows high homology having an  
 CC identity of 88%, with an identity of 81% between the two group I proteins  
 CC (see also AAR51727-30). Fusion peptides may be produced which comprise at  
 CC least two or these antigenic fragments. Each region of these fusion  
 CC peptides may be derived from the same, or different, mite allergens. The  
 CC antigenic fragments may be altered by substitution, deletion or addition  
 CC to enhance their antigenicity. These peptides may be produced by chemical  
 CC synthesis, chemical cleavage of the protein allergen or by recombinant  
 CC techniques. These peptides, or the fusion peptides, when administered to  
 CC a house dust mite sensitive individual, are capable of modifying the  
 CC allergic response of the individual to the allergen. The peptides do not  
 CC bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the  
 CC full length protein allergen. This reduces the major complications of  
 CC standard immunotherapy, which are IgE-mediated responses such as  
 CC anaphylaxis. Exposure of mite allergic patients to these peptides may  
 CC tolerate or anergise appropriate T cell subpopulations such that they  
 CC become unresponsive to mite allergens and do not participate in mounting  
 CC an immune response upon exposure. Administration of the peptides may also  
 CC modify the lymphokine secretion profile as compared with exposure to the  
 CC naturally occurring mite protein allergen  
 XX Sequence 27 AA;  
 SQ

Query Match 90.2%; Score 55; DB 2; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.44;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9  
 |||||  
 Db 14 RMQGGCGSC 22

RESULT 7  
 AAR77161  
 ID AAR77161 standard; peptide; 27 AA.  
 XX  
 AC AAR77161;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 30-MAY-1996 (first entry)  
 XX  
 DE Dermatophagoides pteronyssinus group I peptide DPI-2 (13-39).  
 XX House dust mite; DerpI; DerfI; DerpII; DerfII; allergen; allergy;  
 KW overlapping peptide; screening.  
 XX Dermatophagoides pteronyssinus.  
 OS WO9528424-A1.  
 XX  
 PN 26-OCT-1995.  
 XX  
 PD 12-APR-1995; 95WO-US004481.  
 XX  
 PF 14-APR-1994; 94US-00227772.  
 PR (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PA Chen X, Evans S, Shaked Z, Franzen HM, Kuo M;  
 PI  
 XX

DR WPI; 1995-373765/48.  
 XX  
 PT Compsns. contg. house mite allergen-derived peptide(s), some of which are  
 PT new - are used to treat allergy, and are stable, soluble and able to  
 PT induce T cell non-responsiveness.  
 XX  
 PS Disclosure; Fig 2; 61pp; English.  
 XX  
 CC Claimed therapeutic compositions contain at least one of the peptides DPI  
 CC -21.2 and DPI-22.2 and also at least one of the new peptides DPI-23.31,  
 CC DPI-26.6, DPI-20.9, DPI-22.14 and DPI-25.15. The compositions are  
 CC useful for treating sensitivity to house dust mite allergens. The  
 CC peptides were identified by screening overlapping peptides derived from  
 CC D. pteronyssinus and D. farinae group I and II allergens for T-cell  
 CC reactivity in sensitised individuals. The present sequence is that of  
 CC overlapping peptide DPI-2 (13-39). (Updated on 25-MAR-2003 to correct PR  
 CC field.)  
 XX Sequence 27 AA;  
 SQ  
 Query Match 90.2%; Score 55; DB 2; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.44;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 RMQGGCGSC 9  
 |||||  
 Db 14 RMQGGCGSC 22

## RESULT 8

AAR771913  
 ID AAR771913 standard; peptide; 27 AA.  
 XX  
 AC AAR771913;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 16-DEC-1998 (first entry)  
 XX  
 DE Dermatophagoides Der p I protein peptide DPI-2.  
 XX genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
 KW Der p II; Der f I; Der f II; house dust mite allergy.  
 XX Dermatophagoides.  
 OS  
 XX US5820862-A.  
 PN  
 PD 13-OCT-1998.  
 XX  
 PF 07-JUN-1995; 95US-00482142.  
 XX  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
 PI Shaked Z, Rogers BL;  
 XX  
 DR WPI; 1998-567590/48.  
 XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.  
 PT  
 PS Disclosure; Col 83-84; 155pp; English.  
 XX  
 CC The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen selected  
 CC from the allergens Der p I, Der p II, Der f I, or Der f II. The invention

CC also describes modified peptides having similar or enhanced therapeutic properties as the corresponding, naturally occurring allergen, but having reduced side effects. AAW71912 to AAW72000, and AAW72257 to AAW72330 represent peptides from the present invention. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 27 AA;

Query Match 90.2%; Score 55; DB 2; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.44; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RMQGGCGSC 9  
DB 14 RMQGGCGSC 22

RESULT 9  
ID AAW71970 standard; peptide; 27 AA.

XX AAW71970;

XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 16-DEC-1998 (first entry)

XX Dermatophagoides Der f I protein peptide DFI-2.1.

XX genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
KW Der p II; Der f I; Der f II; house dust mite allergy.

XX Dermatophagoides.

XX US5820862-A.

XX 13-OCT-1998.

XX 07-JUN-1995; 95US-00482142.

XX 14-APR-1993; 93WO-US003471.

XX 14-APR-1994; 94US-00227772.

XX 19-MAY-1995; 95US-00445307.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;

XX Shaked Z, Rogers BL;

XX WPI; 1998-567590/48.

XX Dermatophagoides allergen peptides - useful for treating house dust mite allergy.

XX Disclosure; Col 115-116; 155pp; English.

XX The present invention describes peptides for treating sensitivity to house dust mite allergens from the genus Dermatophagoides. Peptides within the scope of the invention comprise at least one T cell epitope, or preferably at least two T cell epitopes of a protein allergen selected from the allergens Der p I, Der p II, Der f I, or Der f II. The invention also describes modified peptides having similar or enhanced therapeutic properties as the corresponding, naturally occurring allergen, but having reduced side effects. AAW71912 to AAW72000, and AAW72257 to AAW72330 represent peptides from the present invention. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 27 AA;

Query Match 90.2%; Score 55; DB 2; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.44; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RMQGGCGSC 9  
DB 14 RMQGGCGSC 22

RESULT 10

AAVS0361  
ID AAVS0361 standard; peptide; 27 AA.

XX AAVS0361;

XX 25-JAN-2000 (first entry)

XX Dermatophagoides sp major protein allergen DP I-2.

XX Allergen; house dust mite; detection; sensitivity; T cell epitope;  
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
KW Der f I; Der p I; Der p II; Der f II.

XX Dermatophagoides sp.

XX US5968526-A.

XX 19-OCT-1999.

XX 07-JUN-1995; 95US-00478572.

XX 14-APR-1994; 94US-00227772.

XX 12-APR-1995; 95WO-US004481.

XX 19-MAY-1995; 95US-00445307.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;

XX Chen X, Evans S, Kuo M;

XX WPI; 1999-590385/50.

XX Screening individuals for allergic reactions to T cell epitopes of major allergens from house dust mites.

XX Claim 1b; Col 85-86; 158pp; English.

XX This invention describes a novel method (I) for detecting whether an individual is sensitive to Dermatophagoides (house dust mites). The method involves detecting sensitivity to house dust mites in patients, comprising combining a blood sample from the individual with 1 or more isolated T cell epitopes of the protein allergens I and II (DP I) and (DP II) from Dermatophagoides (house dust mites). 32 T cell epitopes with varying, defined amino acids sequences (given in the specification) may be used in (I). The sample and allergens are combined under conditions appropriate for the binding of blood components with the polypeptides. The extent of binding is then indicative of the sensitivity of the patient to house dust mites. (I) may be used to screen individuals for sensitivity to Dermatophagoides (house dust mites). The house dust mite is a major cause of a variety of allergic disorders such as asthma, rhinitis and ectopic dermatitis. AAVS0360-Y50542 and AAVS0546-Y50555 represent house dust mite allergen peptide fragments derived from Der p I, Der f II, Der p I and Der f II

XX Sequence 27 AA;

Query Match 90.2%; Score 55; DB 2; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.44; Mismatches 0; Conservative 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

QY 1 RMQGGCGSC 9  
DB 14 RMQGGCGSC 22

RESULT 11  
AAVS0424

ID AAY50424 standard; peptide; 27 AA.

AC AAY50424;

DT 25-JAN-2000 (first entry)

DE Dermatophagoides sp major protein allergen DF I-2.1.

KW Allergen; house dust mite; detection; sensitivity; T cell epitope;

KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;

KW Der f I; Der p I; Der p II; Der f II.

OS Dermatophagoides sp.

US5968526-A.

PD 19-OCT-1999.

PF 07-JUN-1995; 95US-00478572.

PR 14-APR-1994; 94US-00227772.

PR 12-APR-1995; 95WO-US004481.

PR 19-MAY-1995; 95US-00445307.

PA (IMMU-) IMMULOGIC PHARM CORP.

PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;

PI Chen X, Evans S, Kuo M;

XX WPI; 1999-590385/50.

XX Screening individuals for allergic reactions to T cell epitopes of major

XX allergens from house dust mites.

XX Disclosure; Col 115-116; 158pp; English.

XX This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with I or more  
 CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
 CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the sensitivity  
 CC of the patient to house dust mites. (I) may be used to screen individuals  
 CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
 CC mite is a major cause of a variety of allergic disorders such as asthma,  
 CC rhinitis and ectopic dermatitis. AAY50360-Y50542 and AAY50546-Y50555  
 CC represent house dust mite allergen peptide fragments derived from Der p  
 CC I, Der f II, Der p I and Der f II

XX Sequence 27 AA;

QY Query Match

Best Local Similarity 90.2%; Score 55; DB 2; Length 27;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9

Db 14 RMQGGCGSC 22

RESULT 12

AAU19027

ID AAU19027 standard; peptide; 27 AA.

XX AAU19027;

XX 04-DEC-2001 (first entry)

DE T-cell epitope containing peptide DFI-2.

XX

KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;

KW Der f II; anti-allergenic; immunostimulant; house dust mite allergy;

KW T-cell epitope.

OS Dermatophagoides farinae.

PN US268491-B1.

PD 31-JUL-2001.

PF 07-JUN-1995; 95US-00484296.

PR 16-OCT-1991; 91US-00777859.

PR 08-MAY-1992; 92US-00881396.

PR 14-APR-1993; 93WO-US003471.

PR 14-APR-1994; 94US-00227772.

PR 19-MAY-1995; 95US-00445307.

PA (IMMU-) IMMULOGIC PHARM CORP.

PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;

PI Evans S, Shaked Z;

XX WPI; 2001-549074/61.

XX Peptides comprising T cell groups of the major allergens from

XX Dermatophagoides (house dust mites), useful for treating house dust mite

XX allergy in humans, and for diagnosing sensitivity to house dust mite

XX protein allergens.

XX Claim 5; Fig 4; 158pp; English.

XX The invention relates to an isolated peptide of the major protein

XX allergens of the genus Dermatophagoides, which comprises at least one T

XX cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I

XX or DF II. The isolated peptide comprises at least two regions, each

XX region comprising at least one T cell group of a protein allergen of the

XX genus Dermatophagoides. The regions are derived from the same or

XX different protein allergens of the genus Dermatophagoides. The peptides

XX are useful for treating house dust mite allergy in humans. The peptides

XX are also useful for detecting or diagnosing sensitivity to house dust

XX mite protein allergens. The present peptides have similar or enhanced

XX therapeutic properties as the naturally-occurring allergen, but have

XX reduced side effects, and increased solubility and stability. The present

XX sequence represents an allergenic T-cell epitope containing peptide

XX derived from the Dermatophagoides allergenic proteins

XX Sequence 27 AA;

QY Query Match

Best Local Similarity 90.2%; Score 55; DB 4; Length 27;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSC 9

Db 14 RMQGGCGSC 22

RESULT 13

AAU18964

ID AAU18964 standard; peptide; 27 AA.

XX AAU18964;

XX 04-DEC-2001 (first entry)

DE T-cell epitope containing peptide DFI-2.

XX House dust mite; allergenic peptide; Der p I; Der p II; Der f I;

XX Der f II; anti-allergenic; immunostimulant; house dust mite allergy;

XX T-cell epitope.



OS Dermatophagoides pteronyssinus.  
 XX US6268491-B1.  
 XX 31-JUL-2001.  
 XX PF 07-JUN-1995; 95US-00484296.  
 XX 16-OCT-1991; 91US-00777859.  
 XX 08-MAY-1992; 92US-00881396.  
 XX 14-APR-1993; 93WO-US003471.  
 XX 14-APR-1994; 94US-00237772.  
 XX 19-MAY-1995; 95US-00445307.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 XX Evans S, Shaked Z;  
 XX WPI; 2001-549074/61.  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX Example 3; Fig 3; 158pp; English.  
 XX The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic T-cell epitope containing peptide  
 CC derived from the Dermatophagoides allergenic proteins  
 XX  
 XX SQ Sequence 27 AA;  
 Query Match 90.2%; Score 55; DB 4; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.44; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;  
 QY 1 RMQGGCGSC 9  
 DB |||||  
 14 RMQGGCGSC 22  
 RESULT 14  
 AAR36448  
 ID AAR36448 standard; peptide; 29 AA.  
 XX  
 XX AC AAR36448;  
 XX 25-MAR-2003 (revised)  
 DT 12-AUG-1993 (first entry)  
 XX  
 XX DE DFI-3(21-49), a Dermatophagoides protein allergen.  
 XX T cell epitope; house dust mite; allergy; soluble; Def pI.  
 XX Synthetic.  
 XX WO9308279-A1.  
 XX 29-APR-1993.  
 XX PF 15-OCT-1992; 92WO-US008637.  
 XX 16-OCT-1991; 91US-00777859.  
 XX 08-MAY-1992; 92US-00881396.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
 XX WPI; 1993-152472/18.  
 XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 and treatment of sensitivity to house dust mite.  
 XX Claim 44; Fig 4; 176pp; English.  
 XX The peptide is one of a series of overlapping peptides synthesised by  
 standard techniques to cover the whole Dermatophagoides farinae Def pI  
 sequence. The T cell epitopes of the protein were mapped by detection of  
 CC the peptide's ability to stimulate T cell activity. The peptides may be  
 CC used for diagnosis and treatment of sensitivity to house dust mite  
 CC allergens. When administered to house dust mite sensitive individuals,  
 CC the peptides are capable of modifying the allergic response to the  
 CC allergens. The peptides may be modified for e.g. increasing solubility,  
 CC enhancing therapeutic or preventive efficacy or stability. See also  
 CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX SQ Sequence 29 AA;  
 Query Match 90.2%; Score 55; DB 2; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.46; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;  
 QY 1 RMQGGCGSC 9  
 DB |||||  
 6 RMQGGCGSC 14  
 RESULT 15  
 AAR34688  
 ID AAR34688 standard; peptide; 29 AA.  
 XX  
 XX AC AAR34688;  
 XX 25-MAR-2003 (revised)  
 DT 12-AUG-1993 (first entry)  
 XX  
 XX DE DFI-3(21-49) a Dermatophagoides protein allergen.  
 XX T cell epitope; house dust mite; allergy; soluble; Der pI.  
 XX Synthetic.  
 XX WO9308279-A1.  
 XX 29-APR-1993.  
 XX PF 15-OCT-1992; 92WO-US008637.  
 XX 16-OCT-1991; 91US-00777859.  
 XX 08-MAY-1992; 92US-00881396.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
 XX WPI; 1993-152472/18.  
 XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 and treatment of sensitivity to house dust mite.  
 XX Claim 10; Fig 3; 176pp; English.

XX The peptide is one of a series of overlapping peptides synthesised by  
CC standard techniques to cover the whole Dermatophagoides pteronyssinus Der  
CC PI sequence. The T cell epitopes of the protein were mapped by detection  
CC of the peptide's ability to stimulate T cell activity. The peptides may  
CC be used for diagnosis and treatment of sensitivity to house dust mite  
CC allergens. When administered to house dust mite sensitive individuals,  
CC the peptides are capable of modifying the allergic response to the  
CC allergens. The peptides may be modified for e.g. increasing solubility,  
CC enhancing therapeutic or preventive efficacy or stability. See also  
CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN  
CC field.)

XX  
SQ Sequence 29 AA;

Query Match 90.2%; Score 55; DB 2; Length 29;  
Best Local Similarity 100.0%; Pred. NO. 0.46;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMGGGCGSC 9  
Db 6 RMGGGCGSC 14

Search completed: October 27, 2004, 17:43:57  
Job time : 14.7849 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:50:34 ; Search time 10.3586 Seconds

(without alignments)  
312.993 Million cell updates/sec

Title: US-09-867-159A-3

Perfect score: 61

Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

```
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	10	10	US-09-867-159A-3
2	55	90.2	210	14	US-10-001-245-185
3	55	90.2	211	10	US-09-847-208-95
4	55	90.2	211	14	US-10-001-245-184
5	55	90.2	222	10	US-09-867-159A-2
6	55	90.2	222	14	US-10-001-245-14
7	55	90.2	222	14	US-10-001-245-16
8	55	90.2	222	14	US-10-001-245-18
9	55	90.2	222	14	US-10-001-245-20
10	55	90.2	222	14	US-10-001-245-22
11	55	90.2	222	14	US-10-001-245-24
12	55	90.2	222	14	US-10-001-245-26
13	55	90.2	222	14	US-10-001-245-28

14	55	90.2	222	14	US-10-001-245-30	Sequence 30, Appl
15	55	90.2	222	14	US-10-001-245-32	Sequence 32, Appl
16	55	90.2	222	14	US-10-001-245-34	Sequence 34, Appl
17	55	90.2	222	14	US-10-001-245-88	Sequence 88, Appl
18	55	90.2	246	14	US-10-001-245-181	Sequence 181, Appl
19	55	90.2	320	9	US-09-877-160-2	Sequence 2, Appl
20	55	90.2	320	10	US-09-847-208-79	Sequence 79, Appl
21	55	90.2	320	14	US-10-001-245-179	Sequence 179, Appl
22	55	90.2	321	10	US-09-847-208-73	Sequence 73, Appl
23	55	90.2	321	14	US-10-001-245-180	Sequence 180, Appl
24	55	90.2	321	14	US-10-001-245-183	Sequence 183, Appl
25	55	90.2	327	14	US-10-001-245-182	Sequence 182, Appl
26	46	75.4	376	16	US-10-437-963-148140	Sequence 148140, App
27	45	73.8	225	14	US-10-219-220-162	Sequence 162, App
28	45	73.8	225	16	US-10-437-963-119764	Sequence 119764, App
29	45	73.8	277	14	US-10-219-220-280	Sequence 280, App
30	45	73.8	360	14	US-10-087-714-20	Sequence 20, Appl
31	45	73.8	381	14	US-10-219-220-152	Sequence 152, App
32	45	73.8	390	15	US-10-425-114-52217	Sequence 52217, A
33	45	73.8	436	14	US-10-259-165-108	Sequence 108, App
34	45	73.8	436	14	US-10-259-165-442	Sequence 442, App
35	45	73.8	447	15	US-10-425-114-68471	Sequence 68471, A
36	45	73.8	458	16	US-10-437-963-155525	Sequence 155525, A
37	45	73.8	461	15	US-10-425-114-58026	Sequence 58026, A
38	45	73.8	464	15	US-10-618-644-7	Sequence 7, Appl
39	45	73.8	468	15	US-10-424-599-25638	Sequence 225638, A
40	45	73.8	470	16	US-10-437-963-155526	Sequence 155526, A
41	45	73.8	479	15	US-10-425-114-44438	Sequence 44438, A
42	45	73.8	479	15	US-10-425-114-55418	Sequence 55418, A
43	45	73.8	765	13	US-10-087-192-759	Sequence 759, App
44	43	70.5	88	15	US-10-424-599-145560	Sequence 145560, App
45	43	70.5	89	15	US-10-424-599-166743	Sequence 166743, App

#### ALIGNMENTS

```
RESULT 1
US-09-867-159A-3
; Sequence 3, Application US/03867159A
; Publication No. US20030104013A1
; GENERAL INFORMATION:
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one
; FILE REFERENCE: Bil2812US-antiallis
; CURRENT APPLICATION NUMBER: US/09/867,159A
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: FR01/04370
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: FR01/05929
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PPT
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Comprises epitope from cystine protease.
US-09-867-159A-3
```

Query Match 100.0%; Score 61; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMQGGCGSCN 10

DB 1 RMQGGCGSCN 10



```

; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-18

Query Match          90.2%; Score 55; DB 14; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RMOGGCGSC 9
Db      26 RMOGGCGSC 34

RESULT 9
US-10-001-245-20
; Sequence 20, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-20

Query Match          90.2%; Score 55; DB 14; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RMOGGCGSC 9
Db      26 RMOGGCGSC 34

RESULT 10
US-10-001-245-22
; Sequence 22, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1el mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 222
; TYPE: PRT

```

; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-22

Query Match 90.2%; Score 55; DB 14; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMOGGCGSC 9  
Db 26 RMOGGCGSC 34

## RESULT 11

US-10-001-245-24  
; Sequence 24, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-24

Query Match 90.2%; Score 55; DB 14; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMOGGCGSC 9  
Db 26 RMOGGCGSC 34

## RESULT 12

US-10-001-245-26  
; Sequence 26, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-26

Query Match 90.2%; Score 55; DB 14; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMOGGCGSC 9  
Db 26 RMOGGCGSC 34

## RESULT 13

US-10-001-245-28  
; Sequence 28, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-28

Query Match 90.2%; Score 55; DB 14; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMOGGCGSC 9  
Db 26 RMOGGCGSC 34

## RESULT 14

US-10-001-245-30  
; Sequence 30, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/1H942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-30

Query Match 90.2%; Score 55; DB 14; Length 222;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RMOGGCGSC 9  
Db 26 RMOGGCGSC 34

```

RESULT 15
US-10-001-245-32
; Sequence 32, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/IH942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-32

```

```

Query Match      90.2%; Score 55; DB 14; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RMQGGCGSC 9
Db      26 RMQGGCGSC 34

```

Search completed: October 27, 2004, 18:15:43  
 Job time : 11.3586 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:57 ; Search time 3.54582 Seconds  
(without alignments)  
187.032 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	90.2	27	2	US-08-482-142-10
2	55	90.2	27	2	US-08-482-142-10
3	55	90.2	27	2	US-08-478-572-10
4	55	90.2	27	2	US-08-478-572-73
5	55	90.2	27	3	US-08-484-296-10
6	55	90.2	27	3	US-08-484-296-73
7	55	90.2	27	5	PCT-US95-04481-2
8	55	90.2	29	2	US-08-482-142-11
9	55	90.2	29	2	US-08-482-142-74
10	55	90.2	29	2	US-08-478-572-11
11	55	90.2	29	2	US-08-478-572-74
12	55	90.2	29	3	US-08-484-296-11
13	55	90.2	29	3	US-08-484-296-74
14	55	90.2	29	5	PCT-US95-04481-3
15	55	90.2	222	1	US-07-945-288-11
16	55	90.2	222	1	US-08-462-831-11
17	55	90.2	222	1	US-08-461-809-11
18	55	90.2	222	1	US-08-461-441-11
19	55	90.2	222	5	PCT-US93-08518-11
20	55	90.2	245	1	US-07-945-288-2
21	55	90.2	245	1	US-08-462-831-2
22	55	90.2	245	1	US-08-461-809-2
23	55	90.2	245	1	US-08-461-441-2
24	55	90.2	245	2	US-08-482-142-2
25	55	90.2	245	2	US-08-478-572-2
26	55	90.2	245	3	US-08-460-040-2
27	55	90.2	245	3	US-08-484-296-2

28	55	90.2	245	5	PCT-US93-08518-2	Sequence 2, Appli
29	55	90.2	320	1	US-07-945-288-10	Sequence 10, Appl
30	55	90.2	320	1	US-08-462-831-10	Sequence 10, Appl
31	55	90.2	320	1	US-08-461-809-10	Sequence 10, Appl
32	55	90.2	320	1	US-08-461-441-10	Sequence 10, Appl
33	55	90.2	320	5	PCT-US93-08518-10	Sequence 10, Appl
34	55	90.2	321	1	US-07-945-288-6	Sequence 6, Appli
35	55	90.2	321	1	US-08-462-831-6	Sequence 6, Appli
36	55	90.2	321	1	US-08-461-809-6	Sequence 6, Appli
37	55	90.2	321	1	US-08-461-441-6	Sequence 6, Appli
38	55	90.2	321	2	US-08-482-142-6	Sequence 6, Appli
39	55	90.2	321	2	US-08-478-572-6	Sequence 6, Appli
40	55	90.2	321	3	US-08-484-296-6	Sequence 6, Appli
41	55	90.2	321	5	PCT-US93-08518-6	Sequence 6, Appli
42	45	73.8	225	4	US-09-325-932A-162	Sequence 162, App
43	45	73.8	351	1	US-09-500-651-2	Sequence 2, Appli
44	45	73.8	351	2	US-08-313-591-2	Sequence 2, Appli
45	45	73.8	381	4	US-09-325-932A-152	Sequence 152, App

ALIGNMENTS

RESULT 1  
US-08-482-142-10  
; Sequence 10, Application US/8482142  
; Patent No. 5820862  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-482-142-10

Query Match 90.2%; Score 55; DB 2; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.13; Indels 0;  
 Matches 9; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 RMOGGCGSC 9  
 Db 14 RMOGGCGSC 22

RESULT 2  
 US-08-482-142-73  
 ; Sequence 73, Application US/08482142  
 ; Patent No. 5820862  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/482,142  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/445,307  
 ; FILING DATE: 07 June 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 73:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 27 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-482-142-73

Query Match 90.2%; Score 55; DB 2; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9  
 Db 14 RMOGGCGSC 22

RESULT 3  
 US-08-478-572-10  
 ; Sequence 10, Application US/08478572

; Patent No. 5968526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/478,572  
 ; FILING DATE: 07-June-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 10:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 27 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-478-572-10

Query Match 90.2%; Score 55; DB 2; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9  
 Db 14 RMOGGCGSC 22

RESULT 4  
 US-08-478-572-73  
 ; Sequence 73, Application US/08478572  
 ; Patent No. 5968526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207

```
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
/ STREET: 610 LINCOLN STREET
/ CITY: WALTHAM
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/478,572
/ FILING DATE: 07-June-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION NUMBER: 08/445,307
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CRAIG, ANNE I.
/ REGISTRATION NUMBER: 32,976
/ REFERENCE/DOCKET NUMBER: 017.6US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
/ INFORMATION FOR SEQ ID NO: 73:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 27 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: N-terminal
/ US-08-478-572-73

Query Match          90.2%; Score 55; DB 2; Length 27;
Best Local Similarity 100.0%; Pred.No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9
Db 14 RMQGGCGSC 22

RESULT 5
US-08-484-296-10
/ Sequence 10, Application US/08484296
/ Patent No. 6268491
/ GENERAL INFORMATION:
/ APPLICANT: Garman, Richard
/ APPLICANT: Greenstein, Julia
/ APPLICANT: Kuo, Mei-chang
/ APPLICANT: Rogers, Bruce
/ APPLICANT: Franzen, Henry
/ APPLICANT: Evans, Sean
/ APPLICANT: Shaked, Ze'ev
/ TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
/ TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
/ NUMBER OF SEQUENCES: 207
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
/ STREET: 610 LINCOLN STREET
/ CITY: WALTHAM
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/ CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/08/484,296
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/445,307
/ FILING DATE: 07 June 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CRAIG, ANNE I.
/ REGISTRATION NUMBER: 32,976
/ REFERENCE/DOCKET NUMBER: 017.6US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 27 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: N-terminal
/ US-08-484-296-10

Query Match          90.2%; Score 55; DB 3; Length 27;
Best Local Similarity 100.0%; Pred.No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9
Db 14 RMQGGCGSC 22

RESULT 6
US-08-484-296-73
/ Sequence 73, Application US/08484296
/ Patent No. 6268491
/ GENERAL INFORMATION:
/ APPLICANT: Garman, Richard
/ APPLICANT: Greenstein, Julia
/ APPLICANT: Kuo, Mei-chang
/ APPLICANT: Rogers, Bruce
/ APPLICANT: Franzen, Henry
/ APPLICANT: Chen, Xian
/ APPLICANT: Evans, Sean
/ APPLICANT: Shaked, Ze'ev
/ TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
/ TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
/ NUMBER OF SEQUENCES: 207
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
/ STREET: 610 LINCOLN STREET
/ CITY: WALTHAM
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02154
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/484,296
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/445,307
/ FILING DATE: 07 June 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CRAIG, ANNE I.
/ REGISTRATION NUMBER: 32,976
/ REFERENCE/DOCKET NUMBER: 017.6US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
```

```
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-484-296-73

Query Match          90.2%; Score 55; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9
Db 14 RMOGGCGSC 22

RESULT 7
PCT-US95-04481-2
; Sequence 2, Application PC/TUS9504481
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dust Mit
; NUMBER OF SEQUENCES: 54
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04481
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/227,772
; FILING DATE: April 14, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vanstone, Darlene A.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 017.5 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US95-04481-2

Query Match          90.2%; Score 55; DB 5; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9
Db 14 RMOGGCGSC 22

RESULT 8
US-08-482-142-11
; Sequence 11, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154

; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; FRAGMENT TYPE: N-terminal
US-08-482-142-11

Query Match          90.2%; Score 55; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9
Db 6 RMOGGCGSC 14

RESULT 9
US-08-482-142-74
; Sequence 74, Application US/08482142
; Patent No. 5820862
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-482-142-74

Query Match 90.2% Score 55; DB 2; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.14; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RMQGGGSC 9  
Db 6 RMQGGGSC 14

RESULT 10  
US-08-478-572-11  
Sequence 11, Application US/08478572  
Patent No. 5968526  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal

NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-478-572-11  
Query Match 90.2% Score 55; DB 2; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.14; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RMQGGGSC 9  
Db 6 RMQGGGSC 14  
RESULT 11  
US-08-478-572-74  
Sequence 74, Application US/08478572  
Patent No. 5968526  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal

US-08-478-572-74

Query Match 90.2%; Score 55; DB 2; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9  
 |||||  
 Db 6 RMOGGCGSC 14

## RESULT 12

US-08-484-296-11  
 ; Sequence 11, Application US/08484296  
 ; Patent No. 6268491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,296  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE: 07 June 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 11:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 29 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-484-296-11

Query Match 90.2%; Score 55; DB 3; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9  
 |||||  
 Db 6 RMOGGCGSC 14

## RESULT 13

US-08-484-296-74

; Sequence 74, Application US/08484296  
 ; Patent No. 6268491

; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/484,296  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE: 07 June 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 74:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 29 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-484-296-74

Query Match 90.2%; Score 55; DB 3; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMOGGCGSC 9  
 |||||  
 Db 6 RMOGGCGSC 14

## RESULT 14

PCT-US95-04481-3  
 ; Sequence 3, Application PC/TUS9504481  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dust Mit  
 ; NUMBER OF SEQUENCES: 54  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/04481  
 ; FILING DATE:

```
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/227,772
/ FILING DATE: April 14, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Vanstone, Darlene A.
/ REGISTRATION NUMBER: 35,279
/ REFERENCE/DOCKET NUMBER: 017.5 PCT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 29 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FRAGMENT TYPE: internal
PCT-US95-04481-3

Query Match 90.2%; Score 55; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9
Db 6 RMQGGCGSC 14

RESULT 15
US-07-945-288-11
/ Sequence 11, Application US/07945288
/ Patent No. 5433948
/ GENERAL INFORMATION:
/ APPLICANT: Thomas, Wayne R.
/ APPLICANT: Chuva, Kaw-Yan
/ TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM
/ TITLE OF INVENTION: DERMATOPHAGOIDES (HOUSE DUST MITES)
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: LAHIVE & COCKFIELD
/ STREET: 60 STATE STREET, SUITE 510
/ CITY: BOSTON
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/945,288
/ FILING DATE: 19920910
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 580,555
/ FILING DATE: 11 SEPTEMBER 1990
/ APPLICATION NUMBER: 458,642
/ FILING DATE: 13 FEBRUARY 1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MANDRAGOURAS, AMY E.
/ REGISTRATION NUMBER: P36,207
/ REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 227-7400
/ TELEFAX: (617) 227-5941
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 222 amino acids
/ TYPE: AMINO ACID
/ TOPOLOGY: linear
```

```
/ MOLECULE TYPE: protein
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 50
/ OTHER INFORMATION: /label=Xaa is His or Tyr
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 81
/ OTHER INFORMATION: /label=Xaa is Glu or Lys
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 124
/ OTHER INFORMATION: /label=Xaa is Ala or Val
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 136
/ OTHER INFORMATION: /label=Xaa is Ser or Thr
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 215
/ OTHER INFORMATION: /label=Xaa is Glu or Gln
US-07-945-288-11

Query Match 90.2%; Score 55; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSC 9
Db 26 RMQGGCGSC 34

Search completed: October 27, 2004, 17:53:01
Job time : 4.54582 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:11:24 ; Search time 22.5 seconds  
(without alignments)  
42.763 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR79:\*\*  
1: PIR1:\*\*  
2: PIR2:\*\*  
3: PIR3:\*\*  
4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	34.4	10	2 S70336	napin small chain
2	20	32.8	7	2 A12016	formylglycinamide
3	19	31.1	9	2 I50633	c-rel protein - ch
4	19	31.1	10	2 D61440	polygalacturonase
5	19	31.1	10	2 S27873	hypothetical prote
6	18	29.5	9	1 YFFG	thymic factor - pi
7	18	29.5	9	2 A60957	thymocyte growth p
8	18	29.5	10	2 A61289	streptopain (EC 3
9	17	27.9	3	3 A23751	spinal cord peptid
10	17	27.9	6	2 P0727	T-cell receptor be
11	17	27.9	10	2 S06964	hypothetical prote
12	17	27.9	10	2 A60527	sperm-activating p
13	17	27.9	10	2 H60787	sperm-activating p
14	17	27.9	10	2 F60787	sperm-activating p
15	17	27.9	10	2 A60787	sperm-activating p
16	17	27.9	10	2 D60787	sperm-activating p
17	17	27.9	10	2 B60787	sperm-activating p
18	17	27.9	10	2 D60588	sperm-activating p
19	17	27.9	10	2 B60588	sperm-activating p
20	17	27.9	10	2 C60588	sperm-activating p
21	17	27.9	10	2 I60527	sperm-activating p
22	17	27.9	10	2 A60588	sperm-activating p
23	17	27.9	10	2 A60788	sperm-activating p
24	17	27.9	10	2 D60527	sperm-activating p
25	17	27.9	10	2 C39572	sperm-activating p
26	17	27.9	10	2 F60527	sperm-activating p
27	17	27.9	10	2 C60527	sperm-activating p
28	17	27.9	10	2 E60527	sperm-activating p
29	17	27.9	10	2 G60527	sperm-activating p

30 17 27.9 10 2 E39572 sperm-activating p  
31 17 27.9 10 2 D60788 sperm-activating p  
32 17 27.9 10 2 E60788 sperm-activating p  
33 17 27.9 10 2 C60788 sperm-activating p  
34 17 27.9 10 2 F60589 sperm-activating p  
35 17 27.9 10 2 C60589 sperm-activating p  
36 17 27.9 10 2 D60589 sperm-activating p  
37 17 27.9 10 2 I60588 sperm-activating p  
38 17 27.9 10 2 B60589 sperm-activating p  
39 16 26.2 5 2 F22565 R-phycocerythrin ga  
40 16 26.2 6 4 I79564 hypochemical TGL3  
41 16 26.2 10 2 S51912 hemagglutinin - in  
42 16 26.2 10 2 A35556 hypochemical prote  
43 16 26.2 10 2 B61440 polygalacturonase  
44 15 24.6 4 2 S47552 ubiquitin - rat  
45 15 24.6 5 2 A33882 cadmium-binding pe

ALIGNMENTS

RESULT 1

S70336  
napin small chain s1A - Swedish turnip (fragment)  
C:Species: Brassica napus var. rapifera (Swedish turnip, rutabaga)  
C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 09-Jul-2004  
C:Accession: S70336  
R:Neumann, G.M.; Condron, R.; Thomas, I.; Polya, G.M.  
Biochim. Biophys. Acta 1295, 23-33, 1996  
A:Title: Purification and sequencing of multiple forms of Brassica napus seed napin small  
A:Reference number: S70336; MUID:96283790; PMID:8679670  
A:Accession: S70336  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <NEU>  
A:Cross-references: UNIPROT:Q42469; UNIPROT:Q9S9F2

Query Match 34.4%; Score 21; DB 2; Length 10;  
Best Local Similarity 86.7%; Pred. No. 3.8e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MQGGCG 7  
|||  
Db 3 MQGGSG 8

RESULT 2

A12016  
formylglycinamide ribonucleotide amidotransferase (EC 2.-.-.-) - Chicken (fragment)  
C:Species: Gallus gallus (chicken)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 13-Mar-1997  
C:Accession: A12016; B12016  
R:Ohnoki, S.; Hong, B.S.; Buchanan, J.M.  
Fed. Proc. 35, 1549, 1976  
A:Title: Amino acid sequence at glutamine active site for FGAR-amidotransferase.  
A:Reference number: A91459  
A:Accession: A12016  
A:Molecule type: protein  
A:Residues: 1-7 <OHN>  
A:Experimental source: liver, peptide 1  
A:Accession: B12016  
A:Molecule type: protein  
A:Residues: 1-5 <OH2>  
A:Experimental source: liver, peptide 2  
C:Keywords: transferase

Query Match 32.8%; Score 20; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 GGCGSC 9  
| | | |  
Db 1 GVCDBC 6

## RESULT 3

I50633  
 C-rel protein - chicken (fragment)  
 C:Species: Gallus gallus (chicken)  
 C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
 C:Accession: I50633  
 R:Hannink, M.; Temin, H.M.  
 Oncogene 5, 1843-1850, 1990  
 A:Title: Structure and autoregulation of the c-rel promoter.  
 A:Reference number: I50633; MUID:91133738; PMID:2284104  
 A:Accession: I50633  
 A>Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-9 <HAN>  
 A:Cross-references: UNIPROT:Q92009; EMBL:X56440; NID:G63338; PIDN:CAA39822.1; PID:G58448

Query Match 31.1%; Score 19; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 MQGGG 7  
 : |||  
 DB 3 VSGGAG 8

## RESULT 4

D61440  
 polygalacturonase (EC 3.2.1.15) IV - Aspergillus sp. (fragment)  
 C:Species: Aspergillus sp.  
 C>Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 09-Jul-2004  
 C:Accession: D61440  
 R:Sratilova, E.; Markovic, O.; Skrovinova, D.; Rexova-Benkova, L.; Jorvall, H.  
 J. Protein Chem. 12, 15-22, 1993  
 A:Title: Pectinase Aspergillus sp. polygalacturonase: multiplicity, divergence, and structure.  
 A:Reference number: A61440; MUID:93151962; PMID:8427629  
 A:Accession: D61440  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <STP>  
 A:Cross-references: UNIPROT:Q7M501  
 C:Keywords: Glycosidase; hydrolase; polysaccharide degradation

Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 GSC 9  
 : |||  
 DB 1 GSC 3

## RESULT 5

S27873  
 hypothetical protein 2 LRH-1 5'-region - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Aug-2004  
 C:Accession: S27873  
 R:Tugwood, J.D.; Isseemann, I.; Green, S.  
 submitted to the EMBL Data Library, February 1992  
 A:Description: LRH-1: A nuclear hormone receptor active in the absence of exogenous ligand  
 A:Reference number: S27873  
 A:Accession: S27873  
 A:Molecule type: mRNA  
 A:Residues: 1-10 <TUG>  
 A:Cross-references: UNIPROT:Q61807; EMBL:M81385; NID:G198872; PIDN:AAA39446.1; PID:G1988

Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 SCN 10

DB 3 SCN 5  
 |||

## RESULT 6

YFPG  
 thymic factor - pig  
 N:Alternate names: FTS (facteur thymique serique)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C>Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 09-Jul-2004  
 C:Accession: A01523; A60983  
 R:Pleau, J.M.; Dardenne, M.; Blouquit, Y.; Bach, J.F.  
 J. Biol. Chem. 252, 8045-8047, 1977  
 A:Title: Structural study of circulating thymic factor: a peptide isolated from pig serum  
 A:Reference number: A01523; MUID:78026571; PMID:914862  
 A:Accession: A01523  
 A:Molecule type: protein  
 A:Residues: 1-9 <PLE>  
 A:Cross-references: UNIPROT:P01255  
 R:Bach, J.F.; Dardenne, M.; Pleau, J.M.; Rosa, J.  
 Nature 266, 55-57, 1977  
 A:Title: Biochemical Characterisation of a serum thymic factor.  
 A:Reference number: A60983; MUID:77123829; PMID:300146  
 A:Accession: A60983  
 A:Molecule type: protein  
 A:Residues: 'Z', 2-4, 'Z', 6-9 <BAC>  
 A:Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral in a variety of immunoassays.  
 C:Comment: See PIR:A60957 (sheep) for discussion of another possible N-terminal modification  
 C:Superfamily: thymic factor  
 C:Keywords: pyroglutamic acid  
 P1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

Query Match 29.5%; Score 18; DB 1; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RMQGG 5  
 : |||  
 DB 3 KSQGG 7

## RESULT 7

A60957  
 thymocyte growth peptide - sheep  
 N:Contains: FTS (facteur thymique serique)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A60957  
 R:Ernstroem, U.; Gavvelin, G.; Rudja, J.M.  
 Biosci. Rep. 10, 403-412, 1990  
 A:Title: Purification of thymocyte growth peptide (TGP) from sheep thymus. Relationship to  
 A:Reference number: A60957; MUID:91064427; PMID:2249004  
 A:Accession: A60957  
 A:Molecule type: protein  
 A:Residues: 1-9 <ERN>  
 A:Cross-references: UNIPROT:Q7M3C5  
 A:Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral in a variety of immunoassays.  
 C:Comment: This peptide was isolated in two forms. One form contained the pyroglutamate carboxylic acid (Gln) in thymocyte growth peptide (TGP) from sheep thymus. Relationship to  
 C:Superfamily: thymic factor  
 C:Keywords: blocked amino end; pyroglutamic acid  
 P1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
 P1/Modified site: blocked amino end (Gln) (in thymocyte growth peptide) #status experimental

Query Match 29.5%; Score 18; DB 2; Length 9;  
 Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RMQGG 5  
 : |||  
 DB 3 KSQGG 7

RESULT 8  
A61289  
Streptopain (EC 3.4.22.10) - Streptococcus sp. (ATCC 12112, group A) (fragment)  
C:Species: Streptococcus sp.  
C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 09-Jul-2004  
C:Accession: A61289  
R:Lo, S.S.; Fraser, B.A.; Liu, T.Y.  
J. Biol. Chem. 259, 11041-11045, 1984  
A:Title: The mixed disulfide in the zymogen of streptococcal proteinase. Characterization  
A:Reference number: A61289; MUID:84289536; PMID:6381494  
A:Accession: A61289  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <LOA>  
A:Cross-references: UNIPROT:OTM0M1  
C:Superfamily: streptococcal cysteine proteinase  
C:Keywords: cysteine proteinase; hydrolase

Query Match 29.5%; Score 18; DB 2; Length 10;  
Best Local Similarity 42.9%; Pred. No. 9.8e+03;  
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 QGGCGC 9  
Db 3 QAATGHC 9

RESULT 9  
A23751  
spinal cord peptide SCP-4 - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Mar-2004  
R:Hsi, K.L.; Chen, R.L.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K  
Arch. Biochem. Biophys. 240, 178-183, 1985  
A:Reference number: A23751; MUID:85250425; PMID:4015098  
A:Accession: A23751  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <HSI>

Query Match 27.9%; Score 17; DB 3; Length 3;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QGG 5  
Db 1 QGG 3

RESULT 10  
PT0727  
T-cell receptor beta chain V-D-J region (161-2F) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0727  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0727  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-6 <FEE>  
A:Experimental source: newborn thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 27.9%; Score 17; DB 2; Length 6;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGCG 7  
Db 3 GGLG 6

RESULT 11  
S06964  
hypothetical protein (nifa 5' region) - Rhizobium leguminosarum (fragment)  
C:Species: Rhizobium leguminosarum  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C:Accession: S06964  
R:Roelvink, P.W.; Hontelez, J.G.J.; van Kammen, A.; van den Bos, R.C.  
Mol. Microbiol. 3, 1441-1447, 1989  
A:Title: Nucleotide sequence of the regulatory nifa gene of Rhizobium leguminosarum PRE:  
A:Reference number: S06964; MUID:90136072; PMID:2693897  
A:Accession: S06964  
A:Molecule type: DNA  
A:Residues: 1-10 <ROE>  
A:Cross-references: UNIPROT:Q52837; EMBL:X17073; NID:G46208; PIDN:CAA34923.1; PID:G80974

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 60.0%; Pred. No. 1.4e+04;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QGGCG 7  
Db 1 RGGFG 5

RESULT 12  
A60527  
sperm-activating peptide (Br-Phe-2 SAP-1) - sea urchin (Tripneustes gratilla)  
N:Alternate names: SAP-a; speract; TG-1; TG-9  
C:Species: Tripneustes gratilla  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Aug-2000  
C:Accession: A60527; A39572; A60973; H60527; I39572  
R:Yoshino, K.I.; Kajitara, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamaguchi  
Comp. Biochem. Physiol. B 94, 739-751, 1989  
A:Title: A halogenated amino acid-containing sperm activating peptide and its related peptide  
A:Reference number: A60527  
A:Accession: A60527  
A:Molecule type: protein  
A:Residues: 1-10 <YOS>  
R:Yoshino, K.; Takao, T.; Suhara, M.; Kitai, T.; Hori, H.; Nomura, K.; Yamaguchi, M.; Shi  
Biochemistry 30, 6203-6209, 1991  
A:Title: Identification of a novel amino acid, o-bromo-L-phenylalanine, in egg-associated  
A:Reference number: A39572; MUID:91283461; PMID:2059627  
A:Accession: A39572  
A:Molecule type: protein  
A:Residues: 1-10 <YOS>  
R:Yoshino, K.; Yoshino, K.; Suzuki, N.; Shimonishi, Y.  
Biomed. Environ. Mass Spectrom. 19, 705-712, 1990  
A:Title: Analysis of post-translational modifications of proteins by accurate mass measurement  
A:Reference number: A60973; MUID:91167743; PMID:2076468  
A:Accession: A60973  
A:Molecule type: protein  
A:Residues: 1-10 <TAK>  
C:Superfamily: unassigned animal peptides  
C:Keywords: bromine

Query Match 27.9%; Score 17; DB 2; Length 10;  
Best Local Similarity 75.0%; Pred. No. 1.4e+04;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGCG 7  
Db 7 GGVG 10

RESULT 13  
H60787

sperm-activating peptide (Thr-5 speract) - sea urchin (Anthocidaris crassispina)  
 C/Species: Anthocidaris crassispina  
 C/Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 16-Aug-2004  
 C/Accession: H60787  
 R/Suzuki, N.; Kajitara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
 Comp. Biochem. Physiol. B 89, 687-693, 1988  
 A/Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
 A/Reference number: A60787; PMID:88242184; PMID:3378407  
 A/Accession: H60787  
 A/Molecule type: protein  
 A/Residues: 1-10 <SUZ>  
 A/Cross-references: UNIPROT:Q7M4D6  
 C/Comment: This oligopeptide from egg jelly is one of several from this species, all of  
 at shows some, but not absolute, species restriction.

Query Match 27.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGCG 7  
 |||  
 Db 7 GGVG 10

## RESULT 14

F60787  
 sperm-activating peptide (Ser-5 speract) - sea urchin (Anthocidaris crassispina)  
 C/Species: Anthocidaris crassispina  
 C/Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 16-Aug-2004  
 C/Accession: F60787  
 R/Suzuki, N.; Kajitara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
 Comp. Biochem. Physiol. B 89, 687-693, 1988  
 A/Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
 A/Reference number: A60787; PMID:88242184; PMID:3378407  
 A/Accession: F60787  
 A/Molecule type: protein  
 A/Residues: 1-10 <SUZ>  
 A/Cross-references: UNIPROT:Q7M4D7  
 C/Comment: This oligopeptide from egg jelly is one of several from this species, all of  
 at shows some, but not absolute, species restriction.

Query Match 27.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGCG 7  
 |||  
 Db 7 GGVG 10

## RESULT 15

A60787  
 sperm-activating peptide (speract) - sea urchin (Hemicentrotus pulcherrimus)  
 C/Species: Hemicentrotus pulcherrimus  
 C/Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 09-Jul-2004  
 C/Accession: A60787  
 R/Suzuki, N.; Kajitara, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H.  
 Comp. Biochem. Physiol. B 89, 687-693, 1988  
 A/Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentrotus  
 A/Reference number: A60787; PMID:88242184; PMID:3378407  
 A/Accession: A60787  
 A/Molecule type: protein  
 A/Residues: 1-10 <SUZ>  
 A/Cross-references: UNIPROT:Q25121  
 C/Comment: This oligopeptide from egg jelly is one of several from this species, all of  
 at shows some, but not absolute, species restriction.  
 C/Supfamily: unassigned animal peptides

Query Match 27.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+04;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GGCG 7  
 |||  
 Db 7 GGVG 10

Db 7 GGVG 10  
 |||

Search completed: October 27, 2004, 18:35:27  
 Job time : 23.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:53:09 ; Search time 124.5 Seconds  
(without alignments)  
46.215 Million cell updates/sec

Title: US-09-867-159A-3

Perfect score: 61

Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	47.5	10	2 Q6X2S9	Q6X2S9 homo sapien
2	29	47.5	10	2 AAP83315	Aap83315 homo sapi
3	28	45.9	10	2 Q8WP04	Q8wp04 ateles belz
4	26	42.6	9	2 O12096	O12096 caprine art
5	26	42.6	9	2 O12098	O12098 caprine art
6	26	42.6	9	2 O12100	O12100 caprine art
7	26	42.6	9	2 O12102	O12102 caprine art
8	26	42.6	9	2 O12104	O12104 caprine art
9	21	34.4	10	2 P82938	P82938 hordeum vul
10	20	32.8	10	2 Q9X3M2	Q9x3m2 prochloroco
11	19	31.1	9	2 Q92009	Q92009 gallus gall
12	19	31.1	10	2 Q7M501	Q7m501 aspergillus
13	19	31.1	10	2 Q8NEY9	Q8ney9 homo sapien
14	19	31.1	10	2 Q96QA7	Q96qa7 homo sapien
15	19	31.1	10	2 Q61807	Q61807 mus musculu
16	19	31.1	10	2 Q8UUVW2	Q8uvw2 oreochromis
17	19	31.1	10	2 AAO63472	Aao63472 homo sapi
18	18	29.5	8	2 O15890	O15890 homo sapien
19	18	29.5	9	1 THYFE_PIG	P01255 sus scrofa
20	18	29.5	9	2 Q9H3Y3	Q9h3y3 homo sapien
21	18	29.5	9	2 Q7M3C5	Q7m3c5 ovis aries
22	18	29.5	9	2 Q99JFC	Q99jf4 mus musculu
23	18	29.5	10	2 Q7M0W1	Q7m0w1 streptococc
24	17	27.9	7	1 ASCL_ALLAS	P84071 allium asca
25	17	27.9	8	2 Q9Y4X6	Q9y4x6 homo sapien
26	17	27.9	10	2 Q71VB8	Q71vb8 homo sapien
27	17	27.9	10	2 Q7M3T4	Q7m3t4 tripneustes
28	17	27.9	10	2 Q7M3T5	Q7m3t5 tripneustes
29	17	27.9	10	2 Q7M3T6	Q7m3t6 tripneustes
30	17	27.9	10	2 Q7M3T7	Q7m3t7 tripneustes
31	17	27.9	10	2 Q7M3T8	Q7m3t8 tripneustes

#### ALIGNMENTS

##### RESULT 1

Q6X2S9 PRELIMINARY; PRT; 10 AA.  
ID Q6X2S9;  
AC Q6X2S9;  
DT 05-JUL-2004 (Tremblrel. 27, Created)  
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE Translocase of the inner mitochondrial membrane (Fragment).  
GN Name=TIM23;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Skin;  
RX MEDLINE=22890074; Pubmed=14527731;  
RA Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.;  
RT "Human poly(ADP-ribose) glycohydrolase (PARG) gene and the common promoter sequence it shares with inner mitochondrial membrane  
RT translocase 23 (TIM23).";  
RL Gene 314:181-190(2003).  
DR EMBL; AY258588; AAP83315.1; -.  
FT NON TER 10  
SQ SEQUENCE 10 AA; 893 MW; 2D7A7165B8787878 CRC64;  
Query Match 47.5%; Score 29; DB 2; Length 10;  
Best Local Similarity 56.7%; Pred. No. 6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 MQGGCGSCN 10

Db 1 MEGGGGSGN 9

##### RESULT 2

AAP83315 PRELIMINARY; PRT; 10 AA.  
ID AAP83315;  
AC AAP83315;  
DT 02-MAR-2004 (Tremblrel. 27, Created)  
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)  
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)  
DE Translocase of the inner mitochondrial membrane (Fragment).  
GN TIM23.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Skin;  
RA Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.;  
RT "Human poly(ADP-ribose) glycohydrolase (PARG) gene and the common promoter sequence it shares with inner mitochondrial membrane

```

RT translocase 23 (TM23).";
RL Gene 314:181-190(2003).
DR EMBL; AY258588; AAP83315.1; -.
FT NON TER 10
SQ SEQUENCE 10 AA; 893 MW; 2D7A7165B8787878 CRC64;

Query Match 47.5%; Score 29; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 6e+02; 2; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0;

QY 2 MGGGCGSCN 10
DB 1 MEGGGGSGN 9

RESULT 3
Q8WP04 PRELIMINARY; PRT; 10 AA.
AC Q8WP04;
DT 01-MAR-2002 (T-REMBLrel. 20, Created)
DT 01-MAR-2002 (T-REMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (T-REMBLrel. 20, Last annotation update)
DE Tat protein (Fragment).
GN Name=OAC2;
OS Ateles belzebuth chamek (Chamek spider monkey).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=118643;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21636947; PubMed=11778686;
RA Seanez H.N.; Lima C.R.; Lemos B.; Bonvicino C.R.; Moreira M.A.M.;
RA Canavez F.C.;
FT "Gene assignment in Ateles paniscus chamek (Platyrrhini, Primates).
RT Allocation of 18 markers of human syntenic groups 1,2,7,14,15,17 and
RT 22."
RL Chromosome Res. 9:631-639(2001).
DR EMBL; AF375652; AAL31489.1; -.
FT NON TER 1
SQ SEQUENCE 10 AA; 901 MW; 22DF477DD87EASB8 CRC64;

Query Match 45.9%; Score 28; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.6e+02; 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY 6 CGSC 9
DB 2 CGSC 5

RESULT 4
O12096 PRELIMINARY; PRT; 9 AA.
AC O12096;
DT 01-JUL-1997 (T-REMBLrel. 04, Created)
DT 01-JUL-1997 (T-REMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 19, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Caprine arthritis encephalitis virus (CAEV).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11660;
RN [1]
RP SEQUENCE FROM N.A.
RA Turelli P.; Guiguen F.; Mornex J.-F.; Vigne R.; Querat G.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81439; AAB60832.1; -.
FT NON TER 1
SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.8e+06; 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2;

QY 5 GCGSCN 10
DB 2 GCRLCN 7

RESULT 5
O12098 PRELIMINARY; PRT; 9 AA.
AC O12098;
DT 01-JUL-1997 (T-REMBLrel. 04, Created)
DT 01-JUL-1997 (T-REMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 19, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Caprine arthritis encephalitis virus (CAEV).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11660;
RN [1]
RP SEQUENCE FROM N.A.
RA Turelli P.; Guiguen F.; Mornex J.-F.; Vigne R.; Querat G.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81440; AAB60835.1; -.
FT NON TER 1
SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.8e+06; 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2;

QY 5 GCGSCN 10
DB 2 GCRLCN 7

RESULT 6
O12100 PRELIMINARY; PRT; 9 AA.
AC O12100;
DT 01-JUL-1997 (T-REMBLrel. 04, Created)
DT 01-JUL-1997 (T-REMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 19, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Caprine arthritis encephalitis virus (CAEV).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11660;
RN [1]
RP SEQUENCE FROM N.A.
RA Turelli P.; Guiguen F.; Mornex J.-F.; Vigne R.; Querat G.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81441; AAB60836.1; -.
FT NON TER 1
SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.8e+06; 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 2;

QY 5 GCGSCN 10
DB 2 GCRLCN 7

RESULT 7
O12102 PRELIMINARY; PRT; 9 AA.
AC O12102;
DT 01-JUL-1997 (T-REMBLrel. 04, Created)
DT 01-JUL-1997 (T-REMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (T-REMBLrel. 19, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;

```

OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RP Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; U91442; AAB60838.1; -.  
 DR EMBL; U91442; AAB60838.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 || ||  
 Db 2 GCRLCN 7

## RESULT 8

O12104  
 ID O12104 PRELIMINARY; PRT; 9 AA.  
 AC O12104;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Tat protein (Fragment).  
 GN Namestat;  
 OS Caprine arthritis encephalitis virus (CAEV).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11660;  
 RN [1]  
 RA Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;  
 RP Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; U91443; AAB60840.1; -.  
 DR EMBL; U91443; AAB60840.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 9 AA; 922 MW; 21E8644EB7340EB8 CRC64;

Query Match 42.6%; Score 26; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GCGSCN 10  
 || ||  
 Db 2 GCRLCN 7

## RESULT 9

P82938  
 ID P82938 PRELIMINARY; PRT; 10 AA.  
 AC P82938;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Unknown endosperm protein C (Fragment).  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RA STRAIN=cv. BOMI; TISSUE=Starchy endosperm;  
 RP MEDLINE=21088911; PubMed=11271488;  
 RX Kristoffersen H.E., Flensburg R.;  
 RT "Separation and characterization of basic barley seed proteins.";  
 RL Electrophoresis 21:3693-3700(2000).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 8.5-9.0, ITS MW IS: 11.9 KDA.  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1053 MW; 9E562DC40AA87AAE CRC64;

Query Match 34.4%; Score 21; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGC 6  
 || ||  
 Db 1 GGC 3

RESULT 10  
 Q9X3M2  
 ID Q9X3M2 PRELIMINARY; PRT; 10 AA.  
 AC Q9X3M2;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
 DE Cytochrome b (Fragment).  
 GN Name=petB;  
 OS Prochlorococcus sp.  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 OC Prochlorococcus.  
 OX NCBI\_TaxID=1220;  
 RN [1]  
 RA Urbach E., Chisholm S.W.;  
 RP SEQUENCE FROM N.A.  
 RT "Genetic diversity in Prochlorococcus populations flow cytometrically  
 RT sorted from the Sargasso Sea and Gulf Stream.";  
 RL Limnol. Oceanogr. 43:1615-1630(1998).  
 DR EMBL; AF070219; AAD23269.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1076 MW; 75CA5CB05866D324 CRC64;

Query Match 32.8%; Score 20; DB 2; Length 10;  
 Best Local Similarity 62.5%; Pred. No. 1.5e+04;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RMGGCGS 8  
 || ||  
 Db 3 RKQIGSG 10

## RESULT 11

Q92009  
 ID Q92009 PRELIMINARY; PRT; 9 AA.  
 AC Q92009;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE C-rel protein (P68-c-rel) (Fragment).  
 GN Name=c-rel proto-oncogene;  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=91133738; PubMed=2284104;  
 RX Hannink M., Temin H.M.;  
 RT "Structure and autoregulation of the c-rel promoter.";  
 RL Oncogene 5:1843-1850(1990).  
 RN [2]  
 RA SEQUENCE FROM N.A.  
 RP Hannink M., Temin H.M.;  
 RT "Structure and auto regulation of the c-rel promoter.";  
 RL Oncogene 0:0-0(1990).  
 DR EMBL; X56440; CAA39822.1; -.  
 DR EMBL; X56515; CAA39866.1; -.  
 DR PIR; I50633; I50633.  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 805 MW; DE317DD87865A2CD CRC64;

Query Match 31.1%; Score 19; DB 2; Length 9;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 MQGCG 7  
 Db 3 VSGAG 8

## RESULT 12

Q7M501 Q7M501 PRELIMINARY; PRT; 10 AA.  
 ID Q7M501; 2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyalacturonase (EC 3.2.1.15) IV (Fragment).  
 OS Aspergillus sp.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.  
 OX NCBI\_TaxID=5065;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93151962; PubMed=8427629;  
 RA Stratilova E., Markovic O., Skrovnova D., Rexova-Benkova L.,  
 RA Jorvall H.  
 RT "Pectinase Aspergillus sp. polyalacturonase: multiplicity,  
 RT divergence, and structural patterns linking fungal, bacterial, and  
 RL plant polyalacturonases.";  
 RL J. Protein Chem. 12:15-22(1993).  
 DR PIR; D61440; D61440.  
 DR GO; GO:0004650; F:polyalacturonase activity; IEA.  
 FT NON\_TER 1  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 959 MW; 8452365A1A9DIAE CRC64;

Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GSC 9  
 Db 1 GSC 3

## RESULT 13

Q8NEY9 Q8NEY9 PRELIMINARY; PRT; 10 AA.  
 ID Q8NEY9; 2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Type II hair-specific keratin (Type II hair keratin) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bairwa N.K., Banezai R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bairwa N.K., Banezai R.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY037552; AAK68688.1; -.  
 DR EMBL; AY203963; AAO63472.1; -.  
 DR GO; GO:0005882; C:intermediate filament; IEA.  
 KW Keratin.  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1034 MW; 9B53417EAB458E7E CRC64;

Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CGS 8  
 Db 3 CGS 5

## RESULT 14

Q96QA7 Q96QA7 PRELIMINARY; PRT; 10 AA.  
 ID Q96QA7; 2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE BA151A2.1 (Cdc42 guanine exchange factor (GEF) 9 (Collybistin, PEM-2,  
 DE HPEM-2, KIAA0424)) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Whitehead S.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL451106; CAC88408.1; -.  
 FT NON\_TER 10  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1122 MW; 39925CB878640043 CRC64;

Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+04;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MQGCG 7  
 Db 4 IRGSG 9

## RESULT 15

Q61807 Q61807 PRELIMINARY; PRT; 10 AA.  
 ID Q61807; 1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last annotation update)  
 DE LRH-1 protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu D.L., Liu W.Z., Li Q.L., Wang H.M., Qian D., Treuter E., Zhu C.;  
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M81385; AAA39446.1; -.  
 DR PIR; S27873; S27873.  
 DR MGI; MGI:1346834; N=5a2.  
 SQ SEQUENCE 10 AA; 1133 MW; 998B68F5B7244EA5 CRC64;

Query Match 31.1%; Score 19; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+04;



Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SCN 10

Db 3 SCN 5

Search completed: October 27, 2004, 18:23:47  
Job time : 126.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:51:34 ; Search time 112 Seconds  
(without alignments)  
32.029 Million cell updates/sec

Title: US-09-867-159A-3

Perfect score: 61

Sequence: 1 RVGGGCGSCN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	61	100.0	10	5	Aao20569 Cysteine
2	61	100.0	10	5	Ab598534 Cysteine
3	37	60.7	8	2	Aar77350 Cysteine
4	36	59.0	9	2	Adi11045 Somatosta
5	36	59.0	9	2	Adi11044 Somatosta
6	36	59.0	9	6	Abu55751 Polylinke
7	36	59.0	9	7	Adf69793 M. tuberc
8	36	59.0	10	2	Adi11047 Somatosta
9	32	52.5	10	4	Aag86614 Saccharom
10	31	50.8	8	2	Aaw41566 Human cal
11	31	50.8	8	4	Aab97506 Peptide n
12	30	49.2	6	5	Ab808544 Human HCC
13	30	49.2	6	6	Abu55750 Polylinke
14	30	49.2	6	7	Adf69792 M. tuberc
15	30	49.2	8	4	Aab97503 Peptide n
16	30	49.2	10	2	Aar69298 Gp lib/II
17	30	49.2	10	2	Adb67499 Specific-
18	30	49.2	10	2	Aaw60340 Tumour ho
19	30	49.2	10	2	Aaw50579 GPIIb/III
20	30	49.2	10	2	Aaw93666 Human bre
21	30	49.2	10	2	Aad25476 Tc-99m la
22	30	49.2	10	3	Aay54962 Peptide 1
23	30	49.2	10	3	Aay95503 GPIIb/III
24	30	49.2	10	3	Aab21757 Human bre
25	30	49.2	10	4	Aae06335 Human bre

26	30	49.2	10	5	ABB51909	Abb51909 Human 34P
27	30	49.2	10	5	ABB51815	Abb51815 Human 34P
28	30	49.2	10	5	ABB51593	Abb51593 Human 34P
29	30	49.2	10	5	ABB51491	Abb51491 Human 34P
30	30	49.2	10	5	ABB51705	Abb51705 Human 34P
31	29	47.5	8	2	Aaw52097	Aaw52097 Targettin
32	29	47.5	8	2	Aay27435	Rat-HICP
33	29	47.5	8	5	AAO20971	AAO20971 8-mex ins
34	29	47.5	8	5	AAE15661	AAE15661 Insulin-1
35	29	47.5	8	6	ABP56100	ABP56100 Human IGF
36	29	47.5	8	7	ADH02887	Adh02887 AscII pep
37	29	47.5	8	7	ADH02888	Adh02888 Sali pep
38	29	47.5	9	2	Aaw06390	Cyclic pe
39	29	47.5	9	2	Ad25470	Ad25470 GPIIb/III
40	29	47.5	9	2	Ad25444	Ad25444 GPIIb/III
41	29	47.5	9	2	ADH59029	ADH59029 GPIIb/III
42	29	47.5	9	3	AAy54928	AAy54928 Peptide 1
43	29	47.5	9	4	AAG88625	AAG88625 HER2/NEU
44	29	47.5	9	8	AdE64416	AdE64416 Radiophar
45	29	47.5	9	8	ADH58651	ADH58651 Radiophar

#### ALIGNMENTS

RESULT 1  
AAO20569  
ID AAO20569 standard; peptide; 10 AA.  
XX  
AC AAO20569;  
XX  
XX  
DT 02-JAN-2003 (first entry)  
XX  
DE Cysteine protease epitope peptide region, SEQ ID No 3.  
XX  
KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopic eczema; epitope.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
FN WO200278736-A2.  
XX  
PD 10-OCT-2002.  
XX  
XX  
PF 28-MAR-2002; 2002WO-FR001098.  
XX  
PR 30-MAR-2001; 2001FR-00004370.  
PR 03-MAY-2001; 2001FR-00005929.  
XX 29-MAY-2001; 2001US-00867159.  
XX  
PA (ANTI-) ANTIALIS SARRL.  
XX  
PI Loria E, Terrasse G, Trehin Y;  
XX  
DR WPI; 2002-750636/81.  
XX  
XX  
PT Antiallergic compositions containing an anti-histamine, a histamine  
PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
PT for the allergen.  
XX  
XX  
PS Claim 14; Page 11; 32pp; French.  
XX  
XX  
CC The invention relates to antiallergic compositions containing an anti-  
CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
CC isolated nucleic acid molecule that has at least one polynucleotide  
CC sequence coding for the allergen, together with a pharmaceutical carrier.  
CC The pharmaceutical composition of the invention is useful as a non-  
CC specific antiallergic treatment, and also useful in the treatment of  
CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
CC allergic and atopic eczema. This sequence represents a peptide of a  
CC cysteine protease epitope region relating to the antiallergic

```

CC compositions of the invention
XX
SQ Sequence 10 AA;

Query Match      100.0%; Score 61; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSCN 10
Db 1 RMQGGCGSCN 10

RESULT 2
ABB98534
ID ABB98534 standard; peptide; 10 AA.
XX
AC ABB98534;
XX
DT 13-DEC-2002 (first entry)
XX
DE Cysteine protease epitope #1.
XX
KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunotherapy; allergen; allergic hypersensitivity reaction;
KW allergic asthma; allergic rhinitis; allergic atopic eczema;
KW cysteine protease.
XX
OS Dermatophagoides pteronyssinus.
XX
PN FR2822709-A1.
XX
PD 04-OCT-2002.
XX
PF 03-MAY-2001; 2001FR-00005929.
XX
PR 30-MAR-2001; 2001FR-00004370.
XX
PA (ANTI-) ANTIALIS SARRL.
XX
PI Loria E, Terrasse G, Trehin Y;
XX
DR WPI; 2002-735037/80.
XX
PT Antiallergic composition, useful for preventing and treating e.g. asthma,
PT rhinitis or eczema, containing at least two of allergen, antihistamine
PT and histamine synthesis inhibitor.
XX
PS Claim 8; Page 6; 33pp; French.
XX
CC The present invention relates to an antiallergic pharmaceutical
CC composition (I) comprising a pharmaceutical carrier containing an active
CC agent combination of at least two of: an allergen; an antihistamine; and
CC a histamine synthesis inhibitor. (I) is used for treating or preventing
CC allergic hypersensitivity reactions, especially allergic asthma, allergic
CC rhinitis or allergic atopic eczema, in babies, children or adults. The
CC present sequence is a peptide fragment (epitope) of cysteine protease
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the
CC invention
XX
SQ Sequence 10 AA;

Query Match      100.0%; Score 61; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RMQGGCGSCN 10
Db 1 RMQGGCGSCN 10

RESULT 3
AAR77350

```

---

```

ID AAR77350 standard; peptide; 8 AA.
XX
AC AAR77350;
XX
DT 14-MAY-1996 (first entry)
XX
DE Cysteine proteinase derived peptide #1.
XX
KW Primer; amplify; polymerase chain reaction; PCR; Haemonchus contortus;
KW cysteine proteinase; DM.2; DM.3; DM.4; human; DM.4a; DM.5; antigen;
KW vaccine; helminth; parasite; DM.1; ruminant; gastro-intestinal tract.
XX
OS Haemonchus contortus.
XX
PN W09526402-A1.
XX
PD 05-OCT-1995.
XX
PF 24-MAR-1995; 95WO-GB000665.
XX
PR 25-MAR-1994; 94GB-00005925.
XX
PR 25-MAR-1994; 94GB-00005990.
XX
(MLCW) MALLINCKRODT VETERINARY INC.
PA
PI Knox DP, Smith SK, Smith WD, Redmond D, Murray J;
XX
DR WPI; 1995-351322/45.
DR N-PSDB; AAQ94240.
XX
PT Protective helminth parasite antigen - used in vaccine directed against
PT parasitic nematodes of mammalian gastro-intestinal tract e.g. Haemonchus
PT contortus.
XX
PS Example 16; Fig 15; 79pp; English.
XX
CC The sequences given in AAR77350-53 are peptides derived from the
CC canonical Haemonchus contortus cysteine proteinase molecule which were
CC used in the design of the primers given in AAQ94240-43. These primers
CC were used in the cloning of cDNA fragments from the cysteine proteinase
CC gene, such as DM.1, DM.2, DM.3, DM.4, DM.4a and DM.5 (see also AAQ94246-
CC 51). The amplified fragments may be expressed in a recombinant cell for
CC the production of antigens. These antigens may be used in the preparation
CC of a vaccine against helminth parasites in a human or non-human animal
XX
SQ Sequence 8 AA;

Query Match      60.7%; Score 17; DB 2; Length 8;
Best Local Similarity 85.7%; Pred. No. 1.7e+06;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QGGCGSC 9
Db 1 QGGCGSC 7

RESULT 4
AD111045
ID AD111045 standard; peptide; 9 AA.
XX
AC AD111045;
XX
DT 15-APR-2004 (first entry)
XX
DE Somatostatin analogue peptide SEQ ID NO:60.
XX
KW dihydropyridine/pyridinium salt type redox moiety; retina; lipophilic;
KW antidiabetic; ophthalmological; diabetic retinopathy;
KW growth factor inhibitory activity; somatostatin; somatostatin analogue;
KW octreotide; lanreotide; sequential metabolism;
KW insulin dependent diabetes mellitus.
XX
OS Synthetic.

```

```

XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT Disulfide-bond 4. .7 /note= "Xaa = Nic"
XX FT Modified-site 9 /note= "Xaa = OH"
XX PN WO9912572-A1.
XX XX
XX PD 18-MAR-1999.
XX PF 01-SEP-1998; 98WO-US017987.
XX PR 10-SEP-1997; 97US-0058423P.
XX PA (UYFL ) UNIV FLORIDA.
XX PI Bodor NS, Grant MB;
XX DR WPI; 1999-263364/22.
XX PT New peptide derivatives with redox targetor moiety are useful in the
XX PT treatment of diabetic retinopathy.
XX PS Example 6; SEQ ID NO 60; 186pp; English.
XX CC The present invention describes peptide derivatives (I) comprising a
XX CC dihydropyridine/pyridinium salt type redox moiety for targeting peptides
XX CC to the retina, a bulky lipophilic function and an amino
XX CC acid/dipeptide/tripeptide spacer. Also described: (I) the preparation of
XX CC (I); and (2) intermediate quaternary salts (II). (I) have antidiabetic
XX CC and ophthalmological activities. (I) can be used in the treatment or
XX CC prevention of diabetic retinopathy by delivery of peptides with growth
XX CC factor inhibitory activity (e.g. somatostatin analogues such as
XX CC octreotide and lanreotide) to the retina by sequential metabolism. It is
XX CC envisaged that (I) will be useful in the treatment of insulin dependent
XX CC diabetes mellitus patients for critical periods in diabetic retinopathy
XX CC disease progression before laser photocoagulation is indicated,
XX CC preferably for 1-4 month intervals when a patient is experiencing severe
XX CC non-proliferative diabetic retinopathy or is found to have low risk
XX CC diabetic retinopathy. The present sequence represents a somatostatin
XX CC analogue peptide, which is used in the exemplification of the present
XX CC invention.
XX SQ Sequence 9 AA;
XX Query Match 59.0%; Score 36; DB 2; Length 9;
XX Best Local Similarity 83.3%; Pred. No. 1.7e+06;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 4 GCGGSC 9
XX Db 2 GCGGCG 7
XX RESULT 5
XX ID ADI11044
XX AC ADI11044;
XX DT 15-APR-2004 (first entry)
XX DE Somatostatin analogue peptide SEQ ID NO:59.
XX dihydropyridine/pyridinium salt type redox moiety; retina; lipophilic;
XX antidiabetic; ophthalmological; diabetic retinopathy;
XX growth factor inhibitory activity; somatostatin; somatostatin analogue;
XX octreotide; lanreotide; sequential metabolism;
XX insulin dependent diabetes mellitus.
XX Synthetic.
XX OS

```

```

XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "Xaa = Trig"
XX FT Disulfide-bond 4. .7
XX FT Modified-site 9 /note= "Xaa = OH"
XX PN WO9912572-A1.
XX XX
XX PD 18-MAR-1999.
XX PF 01-SEP-1998; 98WO-US017987.
XX PR 10-SEP-1997; 97US-0058423P.
XX PA (UYFL ) UNIV FLORIDA.
XX PI Bodor NS, Grant MB;
XX DR WPI; 1999-263364/22.
XX PT New peptide derivatives with redox targetor moiety are useful in the
XX PT treatment of diabetic retinopathy.
XX PS Example 6; SEQ ID NO 59; 186pp; English.
XX CC The present invention describes peptide derivatives (I) comprising a
XX CC dihydropyridine/pyridinium salt type redox moiety for targeting peptides
XX CC to the retina, a bulky lipophilic function and an amino
XX CC acid/dipeptide/tripeptide spacer. Also described: (I) the preparation of
XX CC (I); and (2) intermediate quaternary salts (II). (I) have antidiabetic
XX CC and ophthalmological activities. (I) can be used in the treatment or
XX CC prevention of diabetic retinopathy by delivery of peptides with growth
XX CC factor inhibitory activity (e.g. somatostatin analogues such as
XX CC octreotide and lanreotide) to the retina by sequential metabolism. It is
XX CC envisaged that (I) will be useful in the treatment of insulin dependent
XX CC diabetes mellitus patients for critical periods in diabetic retinopathy
XX CC disease progression before laser photocoagulation is indicated,
XX CC preferably for 1-4 month intervals when a patient is experiencing severe
XX CC non-proliferative diabetic retinopathy or is found to have low risk
XX CC diabetic retinopathy. The present sequence represents a somatostatin
XX CC analogue peptide, which is used in the exemplification of the present
XX CC invention.
XX SQ Sequence 9 AA;
XX Query Match 59.0%; Score 36; DB 2; Length 9;
XX Best Local Similarity 83.3%; Pred. No. 1.7e+06;
XX Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 4 GCGGSC 9
XX Db 2 GCGGCG 7
XX RESULT 6
XX ABUS5751
XX ID ABUS5751 standard; peptide; 9 AA.
XX AC ABUS5751;
XX DT 18-MAR-2003 (first entry)
XX DE Polylinker peptide #3 relating to invention of M. tuberculosis antigens.
XX Mycobacterium tuberculosis antigenic polypeptide; immune response;
XX tuberculosis infection; polylinker peptide.
XX Synthetic.
XX OS
XX PN US6465633-B1.
XX XX

```



PT New peptide derivatives with redox targetor moiety are useful in the  
PT treatment of diabetic retinopathy.  
PS Example 10; SEQ ID NO 62; 186pp; English.  
XX  
XX The present invention describes peptide derivatives (I) comprising a  
CC dihydropyridine/pyridinium salt type redox moiety for targeting peptides  
CC to the retina, a bulky lipophilic function and an amino  
CC acid/dipeptide/tripeptide spacer. Also described: (1) the preparation of  
CC (I); and (2) intermediate quaternary salts (II). (I) have antidiabetic  
CC and ophthalmological activities. (II) can be used in the treatment or  
CC prevention of diabetic retinopathy by delivery of peptides with growth  
CC factor inhibitory activity (e.g. somatostatin analogues such as  
CC octreotide and lanreotide) to the retina by sequential metabolism. It is  
CC envisaged that (I) will be useful in the treatment of insulin dependent  
CC diabetes mellitus patients for critical periods in diabetic retinopathy  
CC disease progression before laser photocoagulation is indicated,  
CC preferably for 1-4 month intervals when a patient is experiencing severe  
CC non-proliferative diabetic retinopathy or is found to have low risk  
CC diabetic retinopathy. The present sequence represents a somatostatin  
CC analogue peptide, which is used in the exemplification of the present  
XX invention.  
XX  
SQ Sequence 10 AA;  
Query Match 59.0%; Score 36; DB 2; Length 10;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 4 GCGGSC 9  
| | | | |  
Db 3 GCGGSC 8  
| | | | |  
RESULT 9  
AAG86614  
ID AAG86614 standard; peptide; 10 AA.  
XX  
AC AAG86614;  
XX  
DT 11-SEP-2001 (first entry)  
XX  
DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1563.  
XX  
KW Saccharomyces cerevisiae; complementary peptide; peptide identification;  
KW drug discovery; drug design.  
XX  
XX Saccharomyces cerevisiae.  
OS WO200142276-A1.  
XX  
XX 14-JUN-2001.  
XX  
XX 13-DEC-2000; 2000WO-GB004773.  
XX  
XX 13-DEC-1999; 99GB-00029471.  
XX  
XX (PROT-) PROTEOM LTD.  
XX  
XX Roberts GW, Heal JR;  
XX  
XX WPI; 2001-367863/38.  
DR  
XX  
XX Identifying complementary peptides by analysis of protein and nucleotide  
PT sequence databases, useful in drug design.  
PS Example 3; Page 245; 488pp; English.  
XX  
XX The invention relates to the identification of complementary peptides by  
CC analysis of protein and nucleotide sequence databases from higher  
CC eukaryotic genomes, excluding human and plants. The specific  
CC complementary peptides interact with their relevant target proteins  
CC encoded in the eukaryote genome. The peptides may be used as reagents and

CC drugs for drug discovery and as lead ligands for drug design and  
CC development. The present sequence is a complementary peptide from  
CC Saccharomyces cerevisiae  
XX  
SQ Sequence 10 AA;  
Query Match 52.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 71.4%; Pred. No. 4.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 MQGCGCS 8  
: | | | | |  
Db 4 VSGGCGS 10  
: | | | | |  
RESULT 10  
AAW41566  
ID AAW41566 standard; peptide; 8 AA.  
XX  
AC AAW41566;  
XX  
DT 27-APR-1998 (first entry)  
XX  
DE Human calpain partial peptide.  
XX  
KW Calpain; human; leukocyte; calcium dependent cysteine protease;  
KW screening; activator; inhibitor; treatment; prevention; cancer;  
KW cerebral apoplexy; cerebral infarction; subarachnoid haemorrhage;  
KW Alzheimer's disease; myodystrophy; cataracts; collagen disease;  
KW ischaemic heart disease; atherosclerosis; arthritis.  
XX  
OS Homo sapiens.  
XX  
XX EP799892-A2.  
XX  
PD 08-OCT-1997.  
XX  
PF 03-APR-1997; 97EP-00105508.  
XX  
PR 05-APR-1996; 96JP-00083649.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Shintani Y, Nishi K, Kawamoto T;  
XX  
XX WPI; 1997-482674/45.  
XX  
XX N-PSDB; AAV04204.  
XX  
XX Human calpain protein and related DNA - useful for drug screening and  
PT treating cancer, stroke, etc.  
XX  
XX Disclosure; Page 29; 43pp; English.  
XX  
XX The present sequence is a calpain partial peptide. Calpain is a human  
CC leukocyte derived calcium dependent cysteine protease, which can be used  
CC to screen for compounds that activate or inhibit its proteolytic  
CC activity. Calpain DNA can be used to treat or prevent cancer, cerebral  
CC apoplexy, cerebral infarction, subarachnoid haemorrhage, Alzheimer's  
CC disease, myodystrophy, cataracts, ischaemic heart disease,  
CC atherosclerosis, arthritis or collagen disease  
XX  
XX Sequence 8 AA;  
Query Match 50.8%; Score 31; DB 2; Length 8;  
Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 3 QGCGSC 9  
| | | | |  
Db 2 QGGLGDC 8  
| | | | |  
RESULT 11

AAB97506  
 ID AAB97506 standard; peptide; 8 AA.  
 XX  
 AC AAB97506;  
 XX  
 DT 14-AUG-2001 (first entry)  
 XX  
 DE Peptide nucleic acid peptide fragment #4.  
 XX  
 KW Peptide nucleic acid; PNA; therapy; infection; cancer; restenosis;  
 KW asthma; autoimmune disorder; endocrinological disorder; renal failure;  
 KW neurological disease; acromegaly; sickle cell anaemia;  
 KW polyamide backbone.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "optionally bound to AAB23598"  
 FT Misc-difference 5  
 FT /label= OTHER  
 FT /note= "optionally D-form residue"  
 FT  
 XX US6180767-B1.  
 XX  
 XX 30-JAN-2001.  
 XX  
 XX 07-JAN-1997; 97US-00779072.  
 XX  
 XX 11-JAN-1996; 96US-0009747P.  
 XX  
 XX (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 XX Wickstrom E, Basu S;  
 XX  
 XX WPI; 2001-342005/36.  
 XX  
 XX Novel conjugate, useful for killing pathogenic organisms and for  
 XX inhibiting gene expression, comprising a peptide nucleic acid oligomer  
 XX conjugated to ligand capable of binding to a cell surface receptor via a  
 XX chemical bond or linker.  
 XX  
 XX Example 1; Col 22; 25pp; English.  
 XX  
 XX The present invention provides a number of peptide nucleic acids (PNAs)  
 XX capable of binding to a cell surface receptor, where the oligomer and the  
 XX peptide are linked by a chemical bond or an amino acid linker. The  
 XX oligomer may have a polyamide, polythioamide, polysulfonamide or  
 XX polysulfonamide backbone. The PNAs of the invention can be used in  
 XX therapy, including the treatment of infections, cancer, autoimmune  
 XX diseases, renal failure, endocrinological disorders, acromegaly,  
 XX neurological diseases and sickle cell anaemia. The present sequence is an  
 XX example of a peptide for use in a PNA  
 XX  
 SQ Sequence 8 AA;  
 Query Match 50.8%; Score 31; DB 4; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 GGCGSC 9  
 DB 3 GGCAAC 8  
 |||||  
 RESULT 12  
 ABB08544  
 ID ABB08544 standard; protein; 6 AA.  
 XX  
 XX ABB08544;  
 XX  
 XX 23-MAY-2002 (first entry)  
 XX

XX Human HCCAL PCR primer P4.  
 DE  
 XX HCCAL; liver cancer; cytostatic; human.  
 KW  
 XX Homo sapiens.  
 OS  
 XX CN1322732-A.  
 PN  
 XX 21-NOV-2001.  
 PD  
 XX C8-MAY-2000; 2000CN-00115595.  
 PF  
 XX C8-MAY-2000; 2000CN-00115595.  
 PR  
 XX (SHAN-) SHANGHAI DONGFANG LIVER & GALLBLADDER SU.  
 PA  
 XX Wang H, Zeng J, Wu M;  
 XX  
 XX WPI; 2002-148617/20.  
 DR  
 XX New liver cancer up expressing gene for treating liver cancer and other  
 FT diseases.  
 FT  
 XX Example 5; Page 16 (disclosure); 31pp; Chinese.  
 XX  
 XX The present invention discloses new human HCCAL protein, the  
 CC polynucleotides encoding the polypeptide and the recombinant process to  
 CC produce the polypeptide. The present invention also discloses the method  
 CC of applying the medicine composite of the polypeptide in treating liver  
 CC cancer and other diseases. The present invention also discloses the  
 CC preparation process of HCCAL protein specific antibody and its  
 CC application in diagnosing and treating diseases. The present invention  
 CC also discloses the application of the polynucleotides encoding the new  
 CC HCCAL protein. The present sequence represents a peptide sequence  
 CC relating to human HCCAL protein  
 XX  
 SQ Sequence 6 AA;  
 Query Match 49.2%; Score 30; DB 5; Length 6;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 CGGCN 10  
 DB 1 CGGCN 5  
 |||||  
 RESULT 13  
 ASU55750  
 ID ASU55750 standard; peptide; 6 AA.  
 XX  
 XX AC ASU55750;  
 XX  
 XX 18-MAR-2003 (first entry)  
 DT  
 XX  
 XX Polylinker peptide #2 relating to invention of M. tuberculosis antigens.  
 DE  
 XX Mycobacterium tuberculosis antigenic polypeptide; immune response;  
 KW tuberculosis infection; polylinker peptide.  
 KW  
 XX Synthetic.  
 OS  
 XX US6465633-B1.  
 PN  
 XX 15-OCT-2002.  
 PD  
 XX 23-DEC-1999; 99US-00470191.  
 PF  
 XX 24-DEC-1998; 98US-0113952P.  
 PR  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX



PI Skeiky Y;  
XX WPI; 2003-147072/14.  
XX Novel isolated mycobacterial polynucleotide, useful for treating,  
PT preventing or diagnosing Mycobacterium tuberculosis infection, for  
PT producing Mycobacterium tuberculosis secretory polypeptides and DNA  
PT vaccines.  
XX Disclosure; Col 89; 48pp; English.  
XX The present invention relates to the isolation of polynucleotide  
CC sequences encoding Mycobacterium tuberculosis antigenic polypeptides. The  
CC polynucleotide sequences of the invention are useful for treating,  
CC preventing, and diagnosing M. tuberculosis infection, for producing M.  
CC tuberculosis secretory polypeptides, for producing DNA vaccines, for  
CC diagnostic purposes, as molecular probes or primers to detect the  
CC presence of bacteria in a biological sample, for inducing and/or  
CC enhancing immune responses to M. tuberculosis, and in gene therapy.  
CC ABUS5749-ABUS5754 represent flexible polylinker peptides. Note: The  
CC present sequence is given in the Sequence listing but is not mentioned  
CC elsewhere in the specification  
XX  
XX SQ Sequence 6 AA;  
Query Match 49.2%; Score 30; DB 6; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 GCGGC 9  
DB 1 GCGGC 5  
RESULT 14  
ADF69792  
ID ADF69792 standard; peptide; 6 AA.  
AC ADF69792;  
XX  
XX 12-FEB-2004 (first entry)  
XX M. tuberculosis fusion protein associated flexible polylinker peptide #1.  
XX Fusion protein; Mycobacterium tuberculosis antigen;  
KW tuberculosis infection; immune response; tuberculostatic;  
KW flexible polylinker.  
XX Synthetic.  
XX  
XX US2003147911-A1.  
XX  
XX 07-AUG-2003.  
XX  
XX 05-FEB-2003; 2003US-00359460.  
XX  
XX 13-MAR-1997; 97US-00818112.  
XX 01-OCT-1997; 97US-00942578.  
XX 18-FEB-1998; 98US-00025197.  
XX 07-APR-1998; 98US-00056556.  
XX 30-DEC-1998; 98US-00223040.  
XX 07-APR-1999; 99US-00287849.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Reed SG, Skeiky YA, Dillon DC, Alderson M, Campos-Neto A;  
PI WPI; 2003-897524/82.  
XX  
XX New fusion proteins of Mycobacterium tuberculosis for diagnosing,  
PT preventing or treating tuberculosis infection or in enhancing immune  
PT responses in M. tuberculosis.  
XX

PS Disclosure; SEQ ID NO 42; 135pp; English.  
XX  
XX The present invention relates to fusion proteins of Mycobacterium  
CC tuberculosis antigens, and the polynucleotide sequences encoding them.  
CC The sequences of the invention are useful in a method for preventing  
CC tuberculosis by administering to a subject an amount of the fusion  
CC protein or the polynucleotide that encodes the fusion protein. Also  
CC disclosed is a pharmaceutical composition comprising the fusion protein  
CC or the polynucleotide sequence encoding it. The fusion protein induces an  
CC immune response to M. tuberculosis and can be used in the diagnosis, an  
CC prevention, and treatment of tuberculosis infection. The present sequence  
XX represents a flexible polylinker peptide.  
XX SQ Sequence 6 AA;  
Query Match 49.2%; Score 30; DB 7; Length 6;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 GCGGC 9  
DB 1 GCGGC 5  
RESULT 15  
AAB97503  
ID AAB97503 standard; peptide; 8 AA.  
XX  
XX AAB97503;  
XX  
XX 14-AUG-2001 (first entry)  
XX  
XX Peptide nucleic acid peptide fragment #1.  
XX  
XX Peptide nucleic acid; PNA; therapy; infection; cancer; restenosis;  
KW asthma; autoimmune disorder; endocrinological disorder; renal failure;  
KW neurological disease; acromegaly; sickle cell anaemia;  
KW polyamide backbone.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Modified-site 1 /label= OTHER  
FT /note= "optionally bound to AAB23598"  
FT Modified-site 5 /label= OTHER  
FT /note= "optionally 4-methoxybenzyl-Cys, optionally D-form  
FT residue"  
FT Modified-site 6 /label= OTHER  
FT /note= "optionally modified by benzyl, optionally D-form  
FT residue"  
FT Modified-site 7 /label= OTHER  
FT /note= "optionally modified by phenylmethoxycarbonyl,  
FT optionally D-form residue"  
FT Modified-site 8 /label= OTHER  
FT /note= "optionally modified by phenylmethoxycarbonyl and  
FT resin, optionally D-form residue"  
XX  
XX US6180767-B1.  
XX  
XX 30-JAN-2001.  
XX  
XX 07-JAN-1997; 97US-00779072.  
XX  
XX 11-JAN-1996; 96US-0009747P.  
XX  
XX (UYJE-) UNIV JEFFERSON THOMAS.  
XX Wickstrom E, Basu S;  
PI

XX WPI; 2001-342005/36.  
 XX  
 PT Novel conjugate, useful for killing pathogenic organisms and for  
 PT inhibiting gene expression, comprising a peptide nucleic acid oligomer  
 PT conjugated to ligand capable of binding to a cell surface receptor via a  
 PT chemical bond or linker.  
 XX  
 PS Claim 14; Col 19; 25pp; English.  
 XX  
 CC The present invention provides a number of peptide nucleic acids (PNAs)  
 CC capable of binding to a cell surface receptor, where the oligomer and the  
 CC peptide are linked by a chemical bond or an amino acid linker. The  
 CC oligomer may have a polyamide, polythioamide, polysulfonamide or  
 CC polysulfonamide backbone. The PNAs of the invention can be used in  
 CC therapy, including the treatment of infections, cancer, autoimmune  
 CC diseases, renal failure, endocrinological disorders, acromegaly,  
 CC neurological diseases and sickle cell anaemia. The present sequence is an  
 CC example of a peptide for use in a PNA  
 XX  
 SQ Sequence 8 AA;

Query Match 49.2%; Score 30; DB 4; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GGCGSC 9  
 |||. |  
 Db 3 GGCSKC 8

Search completed: October 27, 2004, 18:19:33  
 Job time : 115 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:19:45 ; Search time 85.5 Seconds  
(without alignments)  
37.920 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMQGGCGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215900 residues  
Total number of hits satisfying chosen parameters: 173299

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NPW\_PUB.pcp.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pcp.\*
  - 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*
  - 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	100.0	10	10	US-09-867-159A-3
2	36	59.0	7	14	Sequence 3, Appl1
3	36	59.0	7	14	Sequence 59, Appl1
4	36	59.0	7	14	Sequence 60, Appl1
5	36	59.0	9	9	Sequence 62, Appl1
6	36	59.0	9	9	Sequence 43, Appl1
7	36	59.0	9	14	Sequence 43, Appl1
8	33	54.1	8	14	Sequence 7, Appl1
9	30	49.2	6	9	Sequence 2, Appl1
10	30	49.2	6	14	Sequence 42, Appl1
11	30	49.2	6	15	Sequence 42, Appl1
12	30	49.2	10	9	Sequence 6, Appl1
13	30	49.2	10	9	Sequence 57, Appl1

14	30	49.2	10	9	US-09-779-308-324	Sequence 324, App
15	30	49.2	10	9	US-09-779-308-436	Sequence 436, App
16	30	49.2	10	9	US-09-779-308-546	Sequence 546, App
17	30	49.2	10	9	US-09-779-308-640	Sequence 640, App
18	30	49.2	10	14	US-10-264-374-57	Sequence 57, Appl1
19	30	49.2	10	14	US-10-375-992-57	Sequence 57, Appl1
20	30	49.2	10	15	US-10-264-374-57	Sequence 57, Appl1
21	30	49.2	10	16	US-10-375-992-57	Sequence 57, Appl1
22	29	47.5	8	13	US-10-010-408-4	Sequence 4, Appl1
23	29	47.5	8	15	US-10-149-138-2467	Sequence 2467, Ap
24	29	47.5	8	15	US-10-149-138-3195	Sequence 3195, Ap
25	29	47.5	8	15	US-10-311-129-26	Sequence 26, Appl1
26	29	47.5	8	16	US-10-149-138-2467	Sequence 2467, Ap
27	29	47.5	8	16	US-10-149-138-3195	Sequence 3195, Ap
28	29	47.5	9	15	US-10-149-138-855	Sequence 855, App
29	29	47.5	9	15	US-10-149-138-2495	Sequence 2495, App
30	29	47.5	9	15	US-10-149-138-3217	Sequence 3217, Ap
31	29	47.5	9	15	US-10-149-138-4122	Sequence 4122, Ap
32	29	47.5	9	16	US-10-149-138-855	Sequence 855, App
33	29	47.5	9	16	US-10-149-138-2495	Sequence 2495, Ap
34	29	47.5	9	16	US-10-149-138-3217	Sequence 3217, Ap
35	29	47.5	9	16	US-10-149-138-4122	Sequence 4122, Ap
36	29	47.5	10	9	US-09-765-086-52	Sequence 52, Appl1
37	29	47.5	10	10	US-09-572-404B-2757	Sequence 2757, Ap
38	29	47.5	10	10	US-09-572-404B-2759	Sequence 2759, Ap
39	29	47.5	10	10	US-09-572-404B-2761	Sequence 2761, Ap
40	29	47.5	10	10	US-09-572-404B-3780	Sequence 3780, Ap
41	29	47.5	10	14	US-10-264-374-52	Sequence 52, Appl1
42	29	47.5	10	14	US-10-375-992-52	Sequence 52, Appl1
43	29	47.5	10	15	US-10-149-138-3100	Sequence 3100, Ap
44	29	47.5	10	15	US-10-149-138-3594	Sequence 3594, Ap
45	29	47.5	10	15	US-10-264-374-52	Sequence 52, Appl1

ALIGNMENTS

RESULT 1  
US-09-867-159A-3  
; Sequence 3, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE REFERENCE: B112812US-antialis and at least one anti-histamine compound  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Comprises epitope from cystine protease.  
US-09-867-159A-3

Query Match 100.0%; Score 61; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.029;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RMQGGCGSCN 10  
Db 1 RMQGGCGSCN 10

RESULT 2  
 US-10-175-833-59  
 ; Sequence 59, Application US/10175833  
 ; Publication No. US20030211981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BODOR, Nicholas Stephen  
 ; APPLICANT: BARTOLOMEO, Maria  
 ; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
 ; FILE REFERENCE: 028724-109  
 ; CURRENT APPLICATION NUMBER: US/10/175,833  
 ; CURRENT FILING DATE: 2002-06-21  
 ; PRIOR APPLICATION NUMBER: US/09/144,991  
 ; PRIOR FILING DATE: 1998-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/058,423  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 59  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; NAME/KEY: BINDING  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Amino acid 1 is attached by Trig.  
 ; FEATURE:  
 ; NAME/KEY: BINDING  
 ; LOCATION: (7)  
 ; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
 ; NAME/KEY: DISULFID  
 ; LOCATION: (3)..(6)  
 ; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
 ; OTHER INFORMATION: attached by a non-peptidal disulfide bond.  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
 US-10-175-833-59

Query Match 59.0%; Score 36; DB 14; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
 |||||  
 Db 1 GCGGSC 6

RESULT 3  
 US-10-175-833-60  
 ; Sequence 60, Application US/10175833  
 ; Publication No. US20030211981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BODOR, Nicholas Stephen  
 ; APPLICANT: BARTOLOMEO, Maria  
 ; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
 ; FILE REFERENCE: 028724-109  
 ; CURRENT APPLICATION NUMBER: US/10/175,833  
 ; CURRENT FILING DATE: 2002-06-21  
 ; PRIOR APPLICATION NUMBER: US/09/144,991  
 ; PRIOR FILING DATE: 1998-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/058,423  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 60  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; NAME/KEY: BINDING  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Amino acid 1 is attached by Nic.

; FEATURE:  
 ; NAME/KEY: BINDING  
 ; LOCATION: (7)  
 ; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
 ; NAME/KEY: DISULFIDE  
 ; LOCATION: (3)..(6)  
 ; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
 ; OTHER INFORMATION: attached by a non-peptidal disulfide bond.  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
 US-10-175-833-60

Query Match 59.0%; Score 36; DB 14; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
 |||||  
 Db 1 GCGGSC 6

RESULT 4  
 US-10-175-833-62  
 ; Sequence 62, Application US/10175833  
 ; Publication No. US20030211981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BODOR, Nicholas Stephen  
 ; APPLICANT: BARTOLOMEO, Maria  
 ; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
 ; FILE REFERENCE: 028724-109  
 ; CURRENT APPLICATION NUMBER: US/10/175,833  
 ; CURRENT FILING DATE: 2002-06-21  
 ; PRIOR APPLICATION NUMBER: US/09/144,991  
 ; PRIOR FILING DATE: 1998-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/058,423  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 62  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; NAME/KEY: BINDING  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Amino acid 1 is attached by Nic.  
 ; FEATURE:  
 ; NAME/KEY: BINDING  
 ; LOCATION: (7)  
 ; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
 ; NAME/KEY: DISULFID  
 ; LOCATION: (3)..(6)  
 ; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
 ; OTHER INFORMATION: attached by a non-peptidal disulfide bond.  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
 US-10-175-833-62

Query Match 59.0%; Score 36; DB 14; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+06;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGSC 9  
 |||||  
 Db 1 GCGGSC 6

RESULT 5  
 US-09-287-849-43  
 ; Sequence 43, Application US/09287849  
 ; Patent No. US20020009459A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-09-287-849-43

Query Match          59.0%; Score 36; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 GCGGSC 9
      |||||
Db      3 GCGGSC 8

RESULT 6
US-10-359-460-43
; Sequence 43, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-09-287-849-43

Query Match          59.0%; Score 36; DB 9; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 GCGGSC 9
      |||||
Db      3 GCGGSC 8

RESULT 6
US-10-359-460-43
; Sequence 43, Application US/10359460
; Publication No. US20030147911A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/10/359,460
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/287,849
; PRIOR FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 9

```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-10-359-460-43

Query Match          59.0%; Score 36; DB 14; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 GCGGSC 9
      |||||
Db      3 GCGGSC 8

RESULT 7
US-10-359-459-7
; Sequence 7, Application US/10359459
; Publication No. US20040013677A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; TITLE OF INVENTION: and Their Uses
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/10/359,459
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US/09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-10-359-459-7

Query Match          59.0%; Score 36; DB 15; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 GCGGSC 9
      |||||
Db      3 GCGGSC 8

RESULT 8
US-10-163-415-2
; Sequence 2, Application US/10163415
; Publication No. US20030129204A1
; GENERAL INFORMATION:
; APPLICANT: KNOX, DAVID PATRICK
; APPLICANT: SMITH, STUART KEVIN
; APPLICANT: SMITH, WILLIAM DAVID
; APPLICANT: REDMOND, DIANE
; APPLICANT: MURRAY, JACQUELINE
; TITLE OF INVENTION: VACCINES AGAINST HELMINTHIC PARASITES
; FILE REFERENCE: 1181-264
; CURRENT APPLICATION NUMBER: US/10/163,415
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: US 08/716418
; PRIOR FILING DATE: 1996-09-20
; PRIOR APPLICATION NUMBER: PCT/GB95/00665
; PRIOR FILING DATE: 1995-03-24
; PRIOR APPLICATION NUMBER: GB 9405925.0
; PRIOR FILING DATE: 1994-03-25
; PRIOR APPLICATION NUMBER: GB 9405990.4

```

; PRIOR FILING DATE: 1994-03-25  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (6)..(6)  
 ; OTHER INFORMATION: The 'Xaa' at location 6 stands for Ser.  
 ; FEATURE:  
 ; OTHER INFORMATION: PCR primer: 508G  
 US-10-163-415-2

Query Match 54.1%; Score 33; DB 14; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QGCGSC 9  
 Db 1 QGCGXC 7

RESULT 9  
 US-09-287-849-42  
 ; Sequence 42, Application US/09287849  
 ; Patent No. US2002009459A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; FILE REFERENCE: 014058-009020US  
 ; CURRENT APPLICATION NUMBER: US/09/287,849  
 ; CURRENT FILING DATE: 1999-04-07  
 ; PRIOR APPLICATION NUMBER: US 08/818,112  
 ; PRIOR FILING DATE: 1997-03-13  
 ; PRIOR APPLICATION NUMBER: US 08/942,578  
 ; PRIOR FILING DATE: 1997-10-01  
 ; PRIOR APPLICATION NUMBER: US 09/025,197  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 09/056,556  
 ; PRIOR FILING DATE: 1998-04-07  
 ; PRIOR APPLICATION NUMBER: US 09/223,040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 42  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:flexible  
 ; OTHER INFORMATION: polylinker  
 US-09-287-849-42

Query Match 49.2%; Score 30; DB 9; Length 6;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GCGSC 9  
 Db 1 GCGGC 5

RESULT 10  
 US-10-359-460-42  
 ; Sequence 42, Application US/10359460

; Publication No. US20030147911A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; FILE REFERENCE: 014058-009020US  
 ; CURRENT APPLICATION NUMBER: US/10/359,460  
 ; CURRENT FILING DATE: 2003-02-05  
 ; PRIOR APPLICATION NUMBER: US/09/287,849  
 ; PRIOR FILING DATE: 1999-04-07  
 ; PRIOR APPLICATION NUMBER: US 08/818,112  
 ; PRIOR FILING DATE: 1997-03-13  
 ; PRIOR APPLICATION NUMBER: US 08/942,578  
 ; PRIOR FILING DATE: 1997-10-01  
 ; PRIOR APPLICATION NUMBER: US 09/025,197  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 09/056,556  
 ; PRIOR FILING DATE: 1998-04-07  
 ; PRIOR APPLICATION NUMBER: US 09/223,040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 42  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:flexible  
 ; OTHER INFORMATION: polylinker  
 US-10-359-460-42

Query Match 49.2%; Score 30; DB 14; Length 6;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GCGSC 9  
 Db 1 GCGGC 5

RESULT 11  
 US-10-359-459-6  
 ; Sequence 6, Application US/10359459  
 ; Publication No. US20040013677A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; FILE REFERENCE: 014058-009010US  
 ; CURRENT APPLICATION NUMBER: US/10/359,459  
 ; CURRENT FILING DATE: 2003-02-05  
 ; PRIOR APPLICATION NUMBER: US/09/223,040  
 ; PRIOR FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:flexible  
 ; OTHER INFORMATION: polylinker  
 US-10-359-459-6

Query Match 49.2%; Score 30; DB 15; Length 6;

Best Local Similarity 80.0%; Pred. No. 1.2e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GCGSC 9  
| | | | |  
Db 1 GCGSC 5

RESULT 12  
US-09-765-086-57  
; Sequence 57, Application US/09765086  
; Patent No. US20010046498A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruuslahti, Erkki  
; APPLICANT: Pasqualini, Renata  
; APPLICANT: Wadlin, Arap  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Ellerby, H. Michael  
; TITLE OF INVENTION: Chimeric Prostate-Homing Peptides With  
; TITLE OF INVENTION: Pro-Apoptotic Activity  
; FILE REFERENCE: P-LJ 3844  
; CURRENT APPLICATION NUMBER: US/09/765,086  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US 09/489,582  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 235  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 57  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide  
US-09-765-086-57

Query Match 49.2%; Score 30; DB 9; Length 10;  
Best Local Similarity 80.0%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CGSCN 10  
| | | | |  
Db 1 CGECN 5

RESULT 13  
US-09-779-308-222  
; Sequence 222, Application US/09779308  
; Patent No. US20020150972A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 34PD7: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER  
; FILE REFERENCE: 129.4USU1  
; CURRENT APPLICATION NUMBER: US/09/779,308  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/181,020  
; PRIOR FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 718  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 222  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-779-308-222

Query Match 49.2%; Score 30; DB 9; Length 10;  
Best Local Similarity 71.4%; Pred. No. 6.8e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMOGGCG 7  
| | | | |  
Db 1 RLQGGAG 7

RESULT 14  
US-09-779-308-324  
; Sequence 324, Application US/09779308  
; Patent No. US20020150972A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 34PD7: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER  
; FILE REFERENCE: 129.4USU1  
; CURRENT APPLICATION NUMBER: US/09/779,308  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/181,020  
; PRIOR FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 718  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 324  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-779-308-324

Query Match 49.2%; Score 30; DB 9; Length 10;  
Best Local Similarity 71.4%; Pred. No. 6.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RMOGGCG 7  
| | | | |  
Db 1 RLQGGAG 7

RESULT 15  
US-09-779-308-436  
; Sequence 436, Application US/09779308  
; Patent No. US20020150972A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Paris  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 34PD7: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER  
; FILE REFERENCE: 129.4USU1  
; CURRENT APPLICATION NUMBER: US/09/779,308  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/181,020  
; PRIOR FILING DATE: 2000-02-08  
; NUMBER OF SEQ ID NOS: 718  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 436  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-779-308-436

Query Match 49.2%; Score 30; DB 9; Length 10;  
Best Local Similarity 71.4%; Pred. No. 6.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RMGGCG 7  
| : | | |  
Db 1 RLGGAG 7

Search completed: October 27, 2004, 18:39:35  
Job time : 88.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:16:05 ; Search time 27.5 Seconds  
(without alignments)  
24.116 Million cell updates/sec

Title: US-09-867-159A-3  
Perfect score: 61  
Sequence: 1 RMQGCGSCN 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/ECTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	59.0	7	4	US-09-144-991B-59
2	36	59.0	7	4	US-09-144-991B-60
3	36	59.0	7	4	US-09-144-991B-62
4	36	59.0	9	4	US-09-470-191-94
5	36	59.0	9	4	US-09-223-040-7
6	36	59.0	9	4	US-09-287-849-43
7	31	50.8	8	2	US-08-835-093A-3
8	31	50.8	8	3	US-09-157-349-3
9	31	50.8	8	3	US-08-779-072A-6
10	30	49.2	6	4	US-09-470-191-93
11	30	49.2	6	4	US-09-223-040-6
12	30	49.2	6	4	US-09-287-849-42
13	30	49.2	8	3	US-08-779-072A-1
14	30	49.2	10	2	US-08-361-864-36
15	30	49.2	10	2	US-08-902-367-7
16	30	49.2	10	3	US-08-535-170-9
17	30	49.2	10	3	US-09-139-802-57
18	30	49.2	10	4	US-09-659-786-57
19	30	49.2	10	4	US-08-926-914-57
20	29	47.5	9	1	US-08-482-880-8
21	29	47.5	9	2	US-08-273-274-8
22	29	47.5	9	2	US-08-475-041-8
23	29	47.5	9	2	US-08-484-773-8
24	29	47.5	10	2	US-08-335-832-8
25	29	47.5	10	3	US-09-141-127-2
26	29	47.5	10	3	US-09-139-802-52
27	29	47.5	10	4	US-09-659-786-52

28	29	47.5	10	4	US-08-926-914-52	Sequence 52, Appl
29	28	45.9	5	1	US-08-467-607-10	Sequence 10, Appl
30	28	45.9	5	2	US-08-469-362-10	Sequence 10, Appl
31	28	45.9	5	2	US-08-850-392-10	Sequence 10, Appl
32	28	45.9	6	4	US-09-982-704-9	Sequence 9, Appl
33	28	45.9	7	3	US-08-827-171B-13	Sequence 13, Appl
34	28	45.9	7	4	US-09-588-995A-111	Sequence 111, Appl
35	28	45.9	7	4	US-09-598-062-13	Sequence 13, Appl
36	28	45.9	8	1	US-08-526-710-28	Sequence 28, Appl
37	28	45.9	8	3	US-08-862-855-28	Sequence 28, Appl
38	28	45.9	8	3	US-09-226-985-28	Sequence 28, Appl
39	28	45.9	8	3	US-09-227-906-28	Sequence 28, Appl
40	28	45.9	8	4	US-09-428-866-28	Sequence 28, Appl
41	28	45.9	9	3	US-08-997-802-10	Sequence 10, Appl
42	28	45.9	9	3	US-08-997-802-11	Sequence 11, Appl
43	28	45.9	10	3	US-09-139-802-32	Sequence 32, Appl
44	28	45.9	10	4	US-09-659-786-32	Sequence 32, Appl
45	28	45.9	10	4	US-08-926-914-32	Sequence 32, Appl

## ALIGNMENTS

RESULT 1  
US-09-144-991B-59  
; Sequence 59, Application US/09144991B  
; Patent No. 6440933  
; GENERAL INFORMATION:  
; APPLICANT: BODOR, Nicholas Stephen  
; APPLICANT: BARTOLOMEO, Maria  
; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
; FILE REFERENCE: 028724-109  
; CURRENT APPLICATION NUMBER: US/09/144,991B  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: US 60/058,423  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 59  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: BINDING  
; LOCATION: (1)  
; OTHER INFORMATION: Amino acid 1 is attached by Trig.  
; NAME/KEY: BINDING  
; LOCATION: (7)  
; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
; NAME/KEY: DISULFID  
; LOCATION: (3)..(6)  
; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
; OTHER INFORMATION: attached by a non-peptidial disulfide bond.  
; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
US-09-144-991B-59

Query Match 59.0%; Score 36; DB 4; Length 7;  
Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
| | | | |  
Db 1 GCGGCG 6

RESULT 2  
US-09-144-991B-60  
; Sequence 60, Application US/09144991B  
; Patent No. 6440933  
; GENERAL INFORMATION:  
; APPLICANT: BODOR, Nicholas Stephen  
; APPLICANT: BARTOLOMEO, Maria

; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
 ; FILE REFERENCE: 028724-109  
 ; CURRENT APPLICATION NUMBER: US/09/144,991B  
 ; PRIOR FILING DATE: 1998-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/058,423  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 60  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; NAME/KEY: BINDING  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Amino acid 1 is attached by Nic.  
 ; NAME/KEY: BINDING  
 ; LOCATION: (7)  
 ; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
 ; NAME/KEY: DISULFIDE  
 ; LOCATION: (3)..(6)  
 ; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
 ; OTHER INFORMATION: attached by a non-peptidic disulfide bond.  
 ; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
 US-09-144-991B-60

Query Match 59.0%; Score 36; DB 4; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
 Db 1 GCGGCG 6

RESULT 3  
 US-09-144-991B-62  
 ; Sequence 62, Application US/09144991B  
 ; Patent No. 6440933  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BODOR, Nicholas Stephen  
 ; APPLICANT: BARTOLOMEO, Maria  
 ; TITLE OF INVENTION: COMPOUNDS AND METHOD FOR THE PREVENTION AND TREATMENT  
 ; FILE REFERENCE: 028724-109  
 ; CURRENT APPLICATION NUMBER: US/09/144,991B  
 ; CURRENT FILING DATE: 1998-09-01  
 ; PRIOR APPLICATION NUMBER: US 60/058,423  
 ; PRIOR FILING DATE: 1997-09-10  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 62  
 ; LENGTH: 7  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; NAME/KEY: BINDING  
 ; LOCATION: (1)  
 ; OTHER INFORMATION: Amino acid 1 is attached by Nic.  
 ; NAME/KEY: BINDING  
 ; LOCATION: (7)  
 ; OTHER INFORMATION: Amino acid 7 is attached by a hydroxide bond.  
 ; NAME/KEY: DISULFID  
 ; LOCATION: (3)..(6)  
 ; OTHER INFORMATION: Cysteine residues at positions 3 and 6 are  
 ; OTHER INFORMATION: attached by a non-peptidic disulfide bond.  
 ; OTHER INFORMATION: Description of Unknown Organism:peptide derivative  
 US-09-144-991B-62

Query Match 59.0%; Score 36; DB 4; Length 7;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
 Db 1 GCGGCG 6

RESULT 4  
 US-09-470-191-94  
 ; Sequence 94, Application US/09470191  
 ; Patent No. 6465633  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Compositions and Methods of Their Use in  
 ; TITLE OF INVENTION: the Treatment, Prevention and Diagnosis of Tuberculosis  
 ; FILE REFERENCE: 014058-008910US  
 ; CURRENT APPLICATION NUMBER: US/09/470,191  
 ; CURRENT FILING DATE: 1999-12-23  
 ; PRIOR APPLICATION NUMBER: US 60/113,952  
 ; PRIOR FILING DATE: 1998-12-24  
 ; NUMBER OF SEQ ID NOS: 97  
 ; SOFTWARE: FastSEQ for Windows Version 3.0  
 ; SEQ ID NO 94  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: flexible polylinker  
 US-09-470-191-94

Query Match 59.0%; Score 36; DB 4; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
 Db 3 GCGGCG 8

RESULT 5  
 US-09-223-040-7  
 ; Sequence 7, Application US/09223040  
 ; Patent No. 6544522  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Alderson, Mark  
 ; APPLICANT: Campos-Neto, Antonio  
 ; APPLICANT: Corixa Corporation  
 ; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens  
 ; TITLE OF INVENTION: and Their Uses  
 ; FILE REFERENCE: 014058-009010US  
 ; CURRENT APPLICATION NUMBER: US/09/223,040  
 ; CURRENT FILING DATE: 1998-12-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:flexible  
 ; OTHER INFORMATION: polylinker  
 US-09-223-040-7

Query Match 59.0%; Score 36; DB 4; Length 9;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+05;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGGSC 9  
 Db 3 GCGGCG 8

## RESULT 6

```

US-09-287-849-43
; Sequence 43, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-09-287-849-43

```

```

Query Match          59.0%; Score 36; DB 4; Length 9;
Best Local Similarity 83.3%; Pred.No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY

4 GCGGCG 9

Db

3 GCGGCG 8

## RESULT 7

```

US-08-835-099A-3
; Sequence 3, Application US/08835099A
; Patent No. 5874277
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,099A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 083649/1996
; FILING DATE: 05-APR-1996
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-835-099A-3

```

```

Query Match          50.8%; Score 31; DB 2; Length 8;
Best Local Similarity 71.4%; Pred.No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY

3 QGCGGC 9

Db

2 QGGLGC 8

## RESULT 8

```

US-09-157-349-3
; Sequence 3, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-157-349-3

Query Match          50.8%; Score 31; DB 3; Length 8;
Best Local Similarity 71.4%; Pred. No. 3.8e+05;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QGCGSC 9
Db 2 QGGLGDC 8

RESULT 9
US-08-779-072A-6
; Sequence 6, Application US/08779072A
; Patent No. 6180767
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric
; APPLICANT: Basu, Sumittra
; TITLE OF INVENTION: PEPTIDE NUCLEIC ACID CONJUGATES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08779,072A
; FILING DATE: January 7, 1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,747
; FILING DATE: January 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Moraco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: NO. 6180767e
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-779-072A-6

Query Match          50.8%; Score 31; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GCGSC 9
Db 3 GGCAAC 8

RESULT 10
US-09-470-191-93
; Sequence 93, Application US/09470191
; Patent No. 6465633

; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods of Their Use in
; FILE REFERENCE: 014058-008910US
; CURRENT APPLICATION NUMBER: US/09/470,191
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/113,952
; FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: flexible polylinker
US-09-470-191-93

Query Match          49.2%; Score 30; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GCGSC 9
Db 1 GCGGC 5

RESULT 11
US-09-223-040-6
; Sequence 6, Application US/09223040
; Patent No. 6544522
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009010US
; CURRENT APPLICATION NUMBER: US/09/223,040
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-09-223-040-6

Query Match          49.2%; Score 30; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GCGSC 9
Db 1 GCGGC 5

RESULT 12
US-09-287-849-42
; Sequence 42, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio

```

```
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 42
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible
; OTHER INFORMATION: polylinker
US-09-287-849-42

Query Match          49.2%; Score 30; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 GCGSC 9
Db      1 GCGGC 5

RESULT 13
US-08-779-072A-1
; Sequence 1, Application US/08779072A
; Patent No. 6180767
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric
; APPLICANT: Basu, Soumitra
; TITLE OF INVENTION: PEPTIDE NUCLEIC ACID CONJUGATES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/779,072A
; FILING DATE: January 7, 1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,747
; FILING DATE: January 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-14
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 6180767e
; INFORMATION FOR SEQ ID NO: 1:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-779-072A-1

Query Match          49.2%; Score 30; DB 3; Length 8;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 GCGSC 9
Db      3 GCGSC 8

RESULT 14
US-08-361-864-36
; Sequence 36, Application US/08361864
; Patent No. 5977064
; GENERAL INFORMATION:
; APPLICANT: Dean, Richard T
; APPLICANT: Lister-James, John
; TITLE OF INVENTION: Multimeric Polyvalent Antithrombotic
; TITLE OF INVENTION: Agents
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,864
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,466A
; FILING DATE: 19921002
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5977064nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,668
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..5
; OTHER INFORMATION: /label= Cyclic
; OTHER INFORMATION: /note= "The sidechain sulfur of the 1st cysteine
; OTHER INFORMATION: is protected by an -CH2CO- group, that also forms
; OTHER INFORMATION: an amide bond with the N-terminus; the Y is the D
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6..8
; OTHER INFORMATION: /label= Tc-99m-binding
; OTHER INFORMATION: /note= "The sidechain sulfur atom of each cysteine
; OTHER INFORMATION: is protected by an acetamido group; the C-terminal
; OTHER INFORMATION: cysteine is an amide"
```

US-08-361-864-36

Query Match 49.2%; Score 30; DB 2; Length 10;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QGGCGSC 9  
 :|||  
 Db 2 RGDCGGC 8

RESULT 15

US-08-902-367-7

; Sequence 7, Application US/08902367

; Patent No. 5997845

; GENERAL INFORMATION:

; APPLICANT: Dean, Richard T.

; APPLICANT: Lister-James, John

; APPLICANT: Civitello, Edgar R.

; TITLE OF INVENTION: Radiolabeled Compounds for Thrombus

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: McDonnell Boehnen Hulbert &amp; Berghoff

; STREET: 300 South Wacker Drive Seventh Floor

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/902,367

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/462,668

; FILING DATE: 05-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5997845nan, Kevin E

; REGISTRATION NUMBER: 35,303

; REFERENCE/DOCKET NUMBER: 90,1104-W

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312 913 0001

; TELEFAX: 312 913 0002

; TELEX:

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1..3

; OTHER INFORMATION: /label= D-Tyr

; OTHER INFORMATION: /note= "The tyrosine residue is in the D-stereo-

; OTHER INFORMATION: chemical configuration"

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1..5

; OTHER INFORMATION: /label= Cyclic

; OTHER INFORMATION: /note= "The sidechain sulfur of the Cys

; OTHER INFORMATION: residue is covalently linked to the amino

; OTHER INFORMATION: terminus by a -CH2CO- group."

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 8..10

; OTHER INFORMATION: /label= Tc-99m-chelator

; OTHER INFORMATION: /note= "The sidechain sulfur atoms of both Cys

; OTHER INFORMATION: residues are each protected with an  
 ; OTHER INFORMATION: acetamidomethyl group"

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 10

; OTHER INFORMATION: /label= Amide

; OTHER INFORMATION: /note= "The carboxyl terminus is modified to an

; OTHER INFORMATION: amide"

US-08-902-367-7

Query Match 49.2%; Score 30; DB 2; Length 10;

Best Local Similarity 57.1%; Pred. No. 1.8e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QGGCGSC 9

:|||

Db 2 RGDCGGC 8

Search completed: October 27, 2004, 18:36:36

Job time : 28.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:12 ; Search time 2.66932 Seconds  
(without alignments)  
360.454 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: Pirl:\*

2: Pirl2:\*

3: Pirl3:\*

4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	100.0	211	2 S21864	probable cysteine
2	55	100.0	245	2 JQ0337	allergen Der p 1 -
3	55	100.0	319	2 A61500	allergen Der f 1 p
4	46	83.6	94	2 S03380	major fecal allerg
5	40	72.7	71	2 E81021	50S ribosomal prot
6	38	69.1	216	2 H69953	conserved hypothet
7	38	69.1	314	2 T18698	hypothetical prote
8	38	69.1	316	2 H97992	hypothetical prote
9	38	69.1	657	2 T12969	hypothetical prote
10	37	67.3	987	2 T23971	hypothetical prote
11	36	65.5	108	2 S69294	hypothetical prote
12	36	65.5	338	2 A81421	tRNA (5-methylamin
13	36	65.5	346	2 I36942	haptoglobin - chim
14	36	65.5	348	1 HPHUR	haptoglobin-relate
15	36	65.5	348	2 AD0704	3-deoxy-D-arabinon
16	36	65.5	367	2 T29599	hypothetical prote
17	36	65.5	440	2 AB1111	B. subtilis YycH p
18	36	65.5	440	2 AF1472	B. subtilis YycH p
19	36	65.5	931	2 D86222	protein F7G19.9 (i
20	35	65.5	1513	2 A70982	probable ATP-depen
21	35	63.6	128	2 D86809	feric uptake regu
22	35	63.6	348	1 ADECH	2-dehydro-3-deoxy-
23	35	63.6	348	2 C90930	hypothetical prote
24	35	63.6	348	2 G85778	hypothetical prote
25	35	63.6	365	2 A30891	regulatory protein
26	35	63.6	380	2 H82856	phospho-2-dehydro-
27	35	63.6	759	2 T46566	anthranilate synth
28	35	63.6	759	2 T39468	anthranilate synth
29	35	63.6	1106	2 S38783	integrin alpha c

30 34 61.8 141 2 A40463 integrin alpha-c  
31 34 61.8 145 2 C41543 integrin alpha-3B  
32 34 61.8 216 2 T12812 hypothetical prote  
33 34 61.8 348 2 I40070 2-dehydro-3-deoxy-  
34 34 61.8 415 2 F89975 aminopeptidase amp  
35 34 61.8 417 2 T33376 hypothetical prote  
36 34 61.8 468 2 S21172 glutamate-tRNA lig  
37 34 61.8 484 1 SYRZET glutamate-tRNA lig  
38 34 61.8 490 2 C70146 glutamate-tRNA lig  
39 34 61.8 525 2 H75514 glutamyl-tRNA synt  
40 34 61.8 536 2 S66716 glutamate-tRNA lig  
41 34 61.8 537 2 D71296 glutamate-tRNA lig  
42 34 61.8 615 2 T15423 hypothetical prote  
43 34 61.8 642 2 C84944 threonine-tRNA lig  
44 34 61.8 859 1 S06418 3',5'-cyclic-GMP p  
45 34 61.8 912 2 T51131 ligand gated chann

#### ALIGNMENTS

##### RESULT 1

S21864

Probable cysteine proteinase (EC 3.4.22.-) - Euroglyphus maynei

N:Alternate names: allergen Eur m 1

C:Species: Euroglyphus maynei

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S21864

R:Kent, N.A.; Hill, M.; Keen, J.N.; Holland, P.W.H.; Hart, B.

submitted to the EMBL Data Library, June 1991

A:Reference number: S21864

A:Accession: S21864

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-211 <KEN>

A:Cross-references: UNIPROT:P25780; EMBL:X60073

C:Genetics:

A:Introns: 100/3; 155/2

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

Query Match 100.0%; Score 55; DB 2; Length 211;

Best Local Similarity 100.0%; Pred. No. 0.0027;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10

DB 167 QPNYHAVNIV 176

##### RESULT 2

JQ0337

allergen Der p 1 - house-dust mite (Dermatophagoides pteronyssinus) (fragment)

C:Species: Dermatophagoides pteronyssinus

C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C:Accession: JQ0337; A27582; A31657; C27634

R:Chua, K.Y.; Stewart, G.A.; Thomas, W.R.; Simpson, R.J.; Plozza, T.M.; J

J Exp. Med. 167, 175-182, 1988

A>Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1:

A:Reference number: JQ0337; MUID:86089411; PMID:3335830

A:Accession: JQ0337

A:Molecule type: mRNA

A:Residues: 1-245 <CHU>

A:Cross-references: UNIPROT:P08176

R:Thomas, W.R.; Stewart, G.A.; Simpson, R.J.; Chua, K.Y.; Plozza, T.M.; Dalworth, R.J.; J

Int. Arch. Allergy Appl. Immunol. 85, 127-129, 1988

A>Title: Cloning and expression of DNA coding for the major house dust mite allergen Der

A:Reference number: A27582; MUID:88114080; PMID:3276629

A:Accession: A27582

A:Molecule type: mRNA

A:Residues: 6-101 <THO>

A:Cross-references: GB:M24794; NID:G387591; PIDN:AAA28296.1; PID:G387592

R:Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.





A:Accession: H69953  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-216 <KUN>  
A:Cross-references: UNIPROT:P54471; GB:Z99116; GB:AL009126; NID:G2634723; PIDN:CAB14448  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yqfN

Query Match 69.1%; Score 38; DB 2; Length 216;  
Best Local Similarity 77.8%; Pred. No. 6.2;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYHAVNI 9  
DB 99 QPNYHAVNI 107

RESULT 7  
T18698  
hypothetical protein B0285.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
R:Stulston, J.  
A:Accession: T18698  
submitted to the EMBL Data Library, June 1994  
A:Reference number: Z19007  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-314 <WIL>  
A:Cross-references: UNIPROT:P46560; EMBL:Z34533; PIDN:CAA84303.1; GSPDB:GN00021; CESP:BO  
A:Experimental source: clone B0285  
C:Genetics:  
A:Gene: CESP.B0285.10  
A:Map position: 3 71/2; 108/2; 163/3; 237/3  
A:Antons: 3/71; 71/2; 108/2; 163/3; 237/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein B0285.10

Query Match 69.1%; Score 38; DB 2; Length 314;  
Best Local Similarity 62.5%; Pred. No. 9.5;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9  
DB 81 PNYHAVNI 88

RESULT 8  
H97992  
hypothetical protein spr0968 [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: H97992  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: H97992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-316 <KUR>  
A:Cross-references: UNIPROT:Q8CYS4; GB:AE007317; PIDN:AAK99772.1; PID:IG15458580; GSPDB:G  
C:Genetics:  
A:Gene: spr0968

Query Match 59.1%; Score 38; DB 2; Length 316;  
Best Local Similarity 56.7%; Pred. No. 9.5;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYHAVNI 9  
DB 250 QPNYHAVNI 258

RESULT 9  
T12969  
hypothetical protein T6H20.180 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T12969  
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sai  
submitted to the Protein Sequence Database, July 1999  
A:Reference number: Z17586  
A:Accession: T12969  
A:Molecule type: DNA  
A:Residues: 1-657 <CHO>  
A:Cross-references: UNIPROT:Q9STF3; EMBL:AL096859; GSPDB:GN00061; ATSP:T6H20.180  
A:Experimental source: cultivar Columbia; BAC clone T6H20  
C:Genetics:  
A:Gene: ATSP:T6H20.180  
A:Map position: 3  
C:Superfamily: Arabidopsis thaliana hypothetical protein At3g41080

Query Match 69.1%; Score 38; DB 2; Length 657;  
Best Local Similarity 60.0%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QPNYHAVNI 10  
DB 11 QPNYHAVNI 20

RESULT 10  
T23971  
hypothetical protein R06C7.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T23971  
R:Gardner, A.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19825  
A:Accession: T23971  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-987 <WIL>  
A:Cross-references: UNIPROT:Q21776; EMBL:Z71266; PIDN:CAA95845.1; GSPDB:GN00019; CESP:RO  
A:Experimental source: clone R06C7  
C:Genetics:  
A:Gene: CESP.R06C7.8  
A:Map position: 1  
A:Antons: 25/2; 202/1; 506/1; 568/3; 825/2

Query Match 67.3%; Score 37; DB 2; Length 987;  
Best Local Similarity 55.6%; Pred. No. 55;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYHAVNI 9  
DB 213 RPNYHGISI 221

RESULT 11  
S69294  
hypothetical protein YLR202c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein J8167.16-a  
C:Species: Saccharomyces cerevisiae  
C>Date: 20-Jul-1996 #sequence\_revision 23-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: S69294  
R:Pauley, A.  
submitted to the EMBL Data Library, September 1994  
A:Description: The sequence of S. cerevisiae cosmid 8167.  
A:Reference number: S48545

Accession: S69294  
 A:Molecule type: DNA  
 A:Residues: 1-108 <PAU>  
 A:Cross-references: UNIPROT:O13531; EMBL:U14913; NID:9544497; PIDN:AA867450.1; PID:92340  
 C:Genetics:  
 A:Gene: MIPS:YLR202c  
 A:Cross-references: SGD:S0004192  
 A:Map position: 12R  
 A:Introns: 21/2  
 C:Superfamily: Saccharomyces hypothetical protein YLR202c

Query Match 65.5%; Score 36; DB 2; Length 108;  
 Best Local Similarity 62.5%; Pred. No. 6.9;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PNYHAVNI 9  
 DB 2 PNFHLINI 9  
 RESULT 12  
 tRNA (5-methylaminomethyl-2-thiouridylyl)-methyltransferase (EC 2.1.1.61) Cj0053c [impr  
 A:Accession: A81421  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C:Accession: A81421  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
 Nature 403, 665-668, 2000  
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: A81421  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-338 <PAR>  
 A:Cross-references: UNIPROT:Q9PJ66; GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB7254  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cj0046; Cj0053c  
 C:Superfamily: probable membrane protein YDL033c  
 C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 65.5%; Score 36; DB 2; Length 338;  
 Best Local Similarity 66.7%; Pred. No. 25;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 QPNYHAVNI 9  
 DB 37 KPNYHEENI 45  
 RESULT 13  
 I36942  
 haptoglobin - chimpanzee (fragment)  
 C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 22-Jun-1999  
 C:Accession: I36942  
 R:Ericsson, L.M.; Kim, H.S.; Maeda, N.  
 Genomics 14, 948-958, 1992  
 A:Title: Junctions between genes in the haptoglobin gene cluster of primates.  
 A:Reference number: I36941; MUID:93122805; PMID:1478675  
 A:Accession: I36942  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-346 <RES>  
 A:Cross-references: GB:M84463; NID:9903708; PIDN:AAA70197.1; PID:9903710  
 C:Genetics:  
 A:Gene: HPR  
 A:Introns: 29/1; 63/1; 88/1  
 C:Superfamily: haptoglobin; complement factor H repeat homology; trypsin homology  
 F;32-85/Domain: complement factor H repeat homology <PH2>  
 F;102-339/Domain: trypsin homology <TRY>

Query Match 65.5%; Score 36; DB 2; Length 346;  
 Best Local Similarity 75.0%; Pred. No. 26;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PNYHAVNI 9  
 DB 180 PNYHQVDI 187  
 RESULT 14  
 HPRUR  
 haptoglobin-related protein precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 09-Jul-2004  
 C:Accession: A00919; A30360; I61855; I60126  
 R:Maeda, N.  
 J. Biol. Chem. 260, 6698-6709, 1985  
 A:Title: Nucleotide sequence of the haptoglobin and haptoglobin-related gene pair. The ha  
 A:Reference number: A92532; MUID:85207676; PMID:2987228  
 A:Accession: A00919  
 A:Molecule type: DNA  
 A:Residues: 1-348 <MAE>  
 A:Cross-references: UNIPROT:P00739; GB:M10935  
 R:Kuhajda, F.P.; Katumuluwa, A.I.; Pasternack, G.R.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 1188-1192, 1989  
 A:Title: Expression of haptoglobin-related protein and its potential role as a tumor anti  
 A:Reference number: A30360; MUID:89145208; PMID:2465547  
 A:Accession: A30360  
 A:Molecule type: protein  
 A:Residues: 68-73,'E',75;92-99,'Q',101-103 <KUH>  
 A:Note: the amino-terminal end of the mature alpha chain was found to be blocked  
 R:Brickson, L.M.; Kim, H.S.; Maeda, N.  
 Genomics 14, 948-958, 1992  
 A:Title: Junctions between genes in the haptoglobin gene cluster of primates.  
 A:Reference number: I36941; MUID:93122805; PMID:1478675  
 A:Accession: I61855  
 A:Status: translation not shown; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-348 <RES>  
 A:Cross-references: GB:M69197; NID:9292156; PIDN:AAA88079.1; PID:g292158  
 R:Maeda, N.; McEvoy, S.M.; Harris, H.F.; Huismann, T.H.; Smithies, O.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 7395-7399, 1986  
 A:Title: Polymorphisms in the human haptoglobin gene cluster: chromosomes with multiple  
 A:Reference number: I60126; MUID:87016954; PMID:2876426  
 A:Accession: I60126  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 258-338,'H',340-348 <RE2>  
 A:Cross-references: GB:M13908; NID:g184320; PIDN:AAA52686.1; PID:g184321  
 A:Experimental source: an individual (C.G.) with multiple haptoglobin-related genes  
 C:Comment: This protein is expressed during pregnancy and in some breast carcinomas. It e  
 C:Genetics:  
 A:Gene: GDB:HPR  
 A:Cross-references: GDB:I19316; OMIM:140210  
 A:Map position: 16q22.1-16q22.1  
 A:Introns: 2/3; 31/1; 65/1; 90/1  
 A:Note: humans are polymorphic in the number of Hpr genes  
 C:Superfamily: haptoglobin; complement factor H repeat homology; trypsin homology  
 F;19/Domain: acute phase; blocked amino end; glycoprotein; hemoglobin binding; heterotetr  
 C:Keywords: signal sequence #status predicted <SIG>  
 F;20-102/Product: haptoglobin-related protein alpha chain #status predicted <ALP>  
 F;34-87/Domain: complement factor H repeat homology <PH2>  
 F;104-348/Product: haptoglobin-related protein beta chain  
 F;104-341/Domain: trypsin homology <TRY>  
 F;20/Modified site: blocked amino end (Leu) (in mature form) #status experimental  
 F;53-87,91-208,251-282,293-323/Disulfide bonds: #status predicted  
 F;126,149,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.5%; Score 36; DB 1; Length 348;  
 Best Local Similarity 75.0%; Pred. No. 26;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 PNYHAVNI 9  
 RESULT 14  
 HPRUR  
 haptoglobin-related protein precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 09-Jul-2004  
 C:Accession: A00919; A30360; I61855; I60126  
 R:Maeda, N.  
 J. Biol. Chem. 260, 6698-6709, 1985  
 A:Title: Nucleotide sequence of the haptoglobin and haptoglobin-related gene pair. The ha  
 A:Reference number: A92532; MUID:85207676; PMID:2987228  
 A:Accession: A00919  
 A:Molecule type: DNA  
 A:Residues: 1-348 <MAE>  
 A:Cross-references: UNIPROT:P00739; GB:M10935  
 R:Kuhajda, F.P.; Katumuluwa, A.I.; Pasternack, G.R.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 1188-1192, 1989  
 A:Title: Expression of haptoglobin-related protein and its potential role as a tumor anti  
 A:Reference number: A30360; MUID:89145208; PMID:2465547  
 A:Accession: A30360  
 A:Molecule type: protein  
 A:Residues: 68-73,'E',75;92-99,'Q',101-103 <KUH>  
 A:Note: the amino-terminal end of the mature alpha chain was found to be blocked  
 R:Brickson, L.M.; Kim, H.S.; Maeda, N.  
 Genomics 14, 948-958, 1992  
 A:Title: Junctions between genes in the haptoglobin gene cluster of primates.  
 A:Reference number: I36941; MUID:93122805; PMID:1478675  
 A:Accession: I61855  
 A:Status: translation not shown; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-348 <RES>  
 A:Cross-references: GB:M69197; NID:9292156; PIDN:AAA88079.1; PID:g292158  
 R:Maeda, N.; McEvoy, S.M.; Harris, H.F.; Huismann, T.H.; Smithies, O.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 7395-7399, 1986  
 A:Title: Polymorphisms in the human haptoglobin gene cluster: chromosomes with multiple  
 A:Reference number: I60126; MUID:87016954; PMID:2876426  
 A:Accession: I60126  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 258-338,'H',340-348 <RE2>  
 A:Cross-references: GB:M13908; NID:g184320; PIDN:AAA52686.1; PID:g184321  
 A:Experimental source: an individual (C.G.) with multiple haptoglobin-related genes  
 C:Comment: This protein is expressed during pregnancy and in some breast carcinomas. It e  
 C:Genetics:  
 A:Gene: GDB:HPR  
 A:Cross-references: GDB:I19316; OMIM:140210  
 A:Map position: 16q22.1-16q22.1  
 A:Introns: 2/3; 31/1; 65/1; 90/1  
 A:Note: humans are polymorphic in the number of Hpr genes  
 C:Superfamily: haptoglobin; complement factor H repeat homology; trypsin homology  
 F;19/Domain: acute phase; blocked amino end; glycoprotein; hemoglobin binding; heterotetr  
 C:Keywords: signal sequence #status predicted <SIG>  
 F;20-102/Product: haptoglobin-related protein alpha chain #status predicted <ALP>  
 F;34-87/Domain: complement factor H repeat homology <PH2>  
 F;104-348/Product: haptoglobin-related protein beta chain  
 F;104-341/Domain: trypsin homology <TRY>  
 F;20/Modified site: blocked amino end (Leu) (in mature form) #status experimental  
 F;53-87,91-208,251-282,293-323/Disulfide bonds: #status predicted  
 F;126,149,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

Db 182 PNYHQVDI 189

RESULT 15  
AD0704  
3-deoxy-D-arabinosephosphonate 7-phosphate synthase [imported] - Salmonella enterica sub  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 22-Jun-2003  
C:Accession: AD0704  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AD0704  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-348 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD02005.1; PID:GI6502843; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY1763  
C:Superfamily: 3-deoxy-7-phosphoheptulonate synthase  
Query Match 65.5%; Score 36; DB 2; Length 348;  
Best Local Similarity 66.7%; Pred. No. 26;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 QPNYHAVNI 9  
Db 237 KPNYHAEDI 245

Search completed: October 27, 2004, 17:51:28  
Job time : 3.66932 secs



OM protein - protein search, using sw model

Run on: October 27, 2004, 17:34:27 ; Search time 14.7809 Seconds  
(without alignments)  
389.270 Million cell updates/sec

Title: US-09-867-159A-4  
Perfect score: 55  
Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	107	2 Q95X04	dermatophag
2	55	100.0	210	2 Q9GY00	dermatophag
3	55	100.0	320	1 MVAL_DERPT	dermatophag
4	55	100.0	321	1 EUMI_BURMA	dermatophag
5	55	100.0	321	1 MVAL_DERFA	dermatophag
6	55	100.0	321	2 BAC53948	dermatophag
7	46	83.6	94	2 Q7M431	dermatophag
8	42	76.4	2876	2 Q9FE41	oryza sativ
9	41	74.5	456	2 Q8XX48	raistonia s
10	40	72.7	71	1 RL31_NEIMA	neisseria m
11	39	70.9	348	1 AROH_ERWHE	erwinia her
12	38	69.1	172	2 Q6XZ06	dermatophag
13	38	69.1	172	2 AAP42768	dermatophag
14	38	69.1	216	1 YQFN_BACSU	dermatophag
15	38	69.1	263	2 Q819P1	dermatophag
16	38	69.1	314	1 KCB3_CABEL	dermatophag
17	38	69.1	316	2 Q8CYS4	dermatophag
18	38	69.1	333	2 Q6QWE9	dermatophag
19	38	69.1	333	2 AAS48506	dermatophag
20	38	69.1	461	2 Q8LNU0	dermatophag
21	38	69.1	497	2 Q6QWFO	dermatophag
22	38	69.1	497	2 Q6GX25	dermatophag
23	38	69.1	497	2 AAS48505	dermatophag
24	38	69.1	515	2 Q6J662	dermatophag
25	38	69.1	515	2 Q6GX23	dermatophag
26	38	69.1	515	2 Q6GX24	dermatophag
27	38	69.1	515	2 AAT10388	dermatophag
28	38	69.1	657	2 Q9STF3	dermatophag
29	37	67.3	73	2 Q6FVN6	dermatophag
30	37	67.3	146	2 Q95X05	dermatophag
31	37	67.3	361	2 Q7VQF7	dermatophag

32	37	67.3	378	2 Q7RPT4	dermatophag
33	37	67.3	474	2 Q6DTW1	dermatophag
34	37	67.3	524	2 Q8KZ20	dermatophag
35	37	67.3	549	2 Q9F6Y1	dermatophag
36	37	67.3	590	2 Q759K4	dermatophag
37	37	67.3	590	2 AAS52193	dermatophag
38	37	67.3	610	2 Q7XAE5	dermatophag
39	37	67.3	925	1 NPA3_MOUSE	dermatophag
40	37	67.3	933	1 NPA3_HUMAN	dermatophag
41	37	67.3	987	2 Q21776	dermatophag
42	37	67.3	1829	2 Q93Y68	dermatophag
43	37	67.3	1923	2 Q7S9W8	dermatophag
44	36	65.5	108	2 Q13531	dermatophag
45	36	65.5	124	2 Q8SZN1	dermatophag

## ALIGNMENTS

RESULT 1					
Q95X04					
ID	Q95X04	PRELIMINARY;		PRT;	107 AA.
AC	Q95X04;				
DT	01-DEC-2001 (TrEMBLrel. 19, Created)				
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	Cysteine proteinase (Fragment).				
GN	Name=CPW3;				
OS	Dermatophagoides farinae (House-dust mite).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Acariiformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;				
OC	Pyroglyphidae; Dermatophagoides.				
OX	NCBI_TaxID=6954;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Park H., Park S.Y., Kim K.Y., Park S.K., Yun H.C.;				
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AF194432; AAL14425.1; -.				
DR	InterPro; IPR000169; Pept_Cys_acsite.				
DR	PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.				
FT	NON_TER 1 107				
FT	NON_TER 107 107				
SQ	SEQUENCE 107 AA; 12277 MW; A80E7876CBA6F97A CRC64;				
Query Match 100.0%; Score 55; DB 2; Length 107;					
Best Local Similarity 100.0%; Pred. No. 0.0066;					
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1 QPNYHAVNIV 10				
Db	90 QPNYHAVNIV 99				
RESULT 2					
Q9GY00					
ID	Q9GY00	PRELIMINARY;		PRT;	210 AA.
AC	Q9GY00;				
DT	01-MAR-2001 (TrEMBLrel. 16, Created)				
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	Allergen Der f1 (Fragment).				
OS	Dermatophagoides farinae (House-dust mite).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Acariiformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;				
OC	Pyroglyphidae; Dermatophagoides.				
OX	NCBI_TaxID=6954;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Hao M.Q., Xu J., Zhong N.S.;				
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.				
CC	-!- SIMILARITY: Belongs to peptidase family C1.				
DR	EMBL; AF285763; AAG00520.1; -.				
DR	PIR; A27634; A27634.				

DR	GO:	GO:0004197; F.cysteine-type endopeptidase activity; IEA.
DR	GO:	GO:0006508; P.proteolysis and peptidolysis; IEA.
DR	InterPro:	IPR000668; Peptidase C1.
DR	InterPro:	IPR000169; Pept_cys_acsite.
DR	Pfam:	PF00112; Peptidase_C1; I.
DR	PRINTS:	PR00705; PAPAIN_1.
DR	SMART:	SM00645; Pept_C1; 1.
DR	PROSITE:	PS00640; THIOL_PROTEASE ASN; 1.
DR	PROSITE:	PS00139; THIOL_PROTEASE CYS; 1.
DR	PROSITE:	PS00639; THIOL_PROTEASE HIS; UNKNOWN_1.
KW	Hydrolase:	Protease; Thiol protease.
FT	NON_TER	1
FT	NON_TER	210
SQ	SEQUENCE	210 AA; 23548 MW; BA08029D642EB90 CRC64;
Query Match 100.0%; Score 55; DB 2; Length 210;		
Best Local Similarity 100.0%; Pred.No. 0.014;		
Matches	10; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	OPNYHAVNIV 10
Ds	166	OPNYHAVNIV 175
RESULT 3		
ID	MMAL_DERPT	STANDARD; PRT; 320 AA.
AC	P08176; Q24616;	
DT	01-AUG-1988 (Rel. 08, Created)	
DT	01-FEB-1995 (Rel. 31, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Major mite fecal allergen Der p 1 precursor (EC 3.4.22.-) (Der p I).	
GN	Names=DERP1,	
OS	Dermatophagoides pteronyssinus (House-dust mite).	
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;	
OC	Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;	
OC	Pyroglyphidae; Dermatophagoides.	
OX	NCBI_TaxID=6956;	
[1]		
RN	SEQUENCE FROM N.A., AND POLYMORPHISM.	
RX	MEDLINE=93357682; PubMed=8353459;	
RX	Chua K.Y., Kehal P.K., Thomas W.R.;	
RT	"Sequence polymorphisms of cDNA clones encoding the mite allergen Der p I.";	
RL	Int. Arch. Allergy Immunol. 101:364-368(1993).	
RN	[2]	
RN	SEQUENCE OF 76-320 FROM N.A.	
RX	MEDLINE=88089411; PubMed=3335830;	
RX	Chua K.Y., Stewart G.A., Thomas W.R., Simpson R.J., Dilworth R.J.,	
RA	Plozza T.M., Turner K.J.;	
RA	"Sequence analysis of cDNA coding for a major house dust mite	
RT	allergen, Der p 1. Homology with cysteine proteases.";	
FL	J. Exp. Med. 167:175-182(1988).	
RN	[3]	
RN	SEQUENCE OF 81-176 FROM N.A.	
RX	MEDLINE=88114080; PubMed=3276629;	
RA	Thomas W.R., Stewart G.A., Simpson R.J., Chua K.Y., Plozza T.M.,	
RA	Dilworth R.J., Nisbet A., Turner K.J.;	
RT	"Cloning and expression of a DNA coding for the major house dust mite	
RT	allergen Der p 1 in Escherichia coli.";	
INT	Arch. Allergy Appl. Immunol. 85:127-129(1988).	
RN	[4]	
RN	REVISIONS TO 232-241.	
RP	MEDLINE=91215493; PubMed=2021874;	
RA	Dilworth R.J., Chua K.Y., Thomas W.R.;	
RT	"Sequence analysis of cDNA coding for a major house dust mite	
RT	allergen, Der p I.";	
FL	Clin. Exp. Allergy 21:25-32(1991).	
RN	[5]	
RN	SEQUENCE OF 99-308 FROM N.A.	
RX	MEDLINE=93130112; PubMed=1483062;	
RA	Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;	
RT	"Molecular characterisation of group I allergen Eur m I from house	

```

FT VARIANT 179 179 E -> K.
FT VARIANT 222 222 V -> A.
FT VARIANT 234 234 S -> T.
FT VARIANT 313 313 E -> Q.
SQ SEQUENCE 320 AA; 36104 MW; A0B1F4DD09791DFE CRC64;

Query Match 100.0%; Score 55; DB 1; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 264 QPNYHAVNIV 273

RESULT 4
EUMI EURMA
ID EUMI EURMA STANDARD; PRT; 321 AA.
AC P25780; Q9TZ23; Q9TZ24; Q9UBA0;
DT 01-MAY-1992 (Rel. 22, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Mite group 1 allergen Eur m 1 precursor (EC 3.4.22.-) (Eur m I).
GN Name=EURM1;
OS Euroglyphus maynei (Mayne's house dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Euroglyphus.
OX NCBI_TaxID=6958;
RN [1]
RP SEQUENCE FROM N.A. (EUR M 1.0101 AND EUR M 1.0102).
RX MEDLINE=99126275; PubMed=9925958;
RA Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;
RT "Molecular analysis of the group 1 and 2 allergens from the house dust mite, Euroglyphus maynei.";
RL Int. Arch. Allergy Immunol. 118:15-22 (1999).
RN [2]
RP SEQUENCE OF 99-309 FROM N.A.
RX MEDLINE=93130112; PubMed=1483062;
RA Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;
RT "Molecular characterisation of group 1 allergen Eur m I from house dust mite Euroglyphus maynei.";
RL Int. Arch. Allergy Immunol. 99:150-152 (1992).
CC -!- FUNCTION: Probable thiol protease.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of mite allergy are bronchial asthma, allergic rhinitis and conjunctivitis.
CC -!- SIMILARITY: Belongs to peptidase family C1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
DR EMBL; AF047610; AAC82351.1; -
DR EMBL; AF047611; AAC82352.1; ALT_INIT.
DR EMBL; AF047612; AAC82353.1; -
DR EMBL; X60073; CA42677.1; -
DR FIR; S21864; S21864.
DR HSSP; PS3634; IK3B.
DR MEROPS; C01.073; -.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Peptidase_C1.
DR Pfam; PF00112; Peptidase_C1; I.
DR PRINTS; PR00705; PAPAIN_C1; I.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Peptidase_C1; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.

```

```

DR PROSITE; PS00639; THIOI_PROTEASE_HIS; 1.
KW Allergen; Glycoprotein; Hydrolase; Signal; Thiol protease; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 98
FT CHAIN 99 321 Mite group 1 allergen Eur m 1.
FT ACT_SITE 133 133 By similarity.
FT ACT_SITE 269 269 By similarity.
FT ACT_SITE 289 289 By similarity.
FT DISULFID 130 170 By similarity.
FT CARBOHYD 34 34 N-linked (GLCNAC...) (Potential).
FT CARBOHYD 151 151 T -> S (in Eur m 1.0102).
FT VARIANT 36 36 M -> N (in Eur m 1.0102).
FT VARIANT 126 126 M -> I (in Eur m 1.0102).
FT VARIANT 320 320 M -> I (in Eur m 1.0102).
SQ SEQUENCE 321 AA; 36290 MW; 6CFD44FEC725999E CRC64;

Query Match 100.0%; Score 55; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 265 QPNYHAVNIV 274

RESULT 5
MVAL DERFA
ID MVAL DERFA STANDARD; PRT; 321 AA.
AC P46311;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Major mite fecal allergen Der f 1 precursor (EC 3.4.22.-) (Der f I).
GN Name=DERf1;
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9215493; PubMed=2021874;
RA Dillworth R.J., Chua K.Y., Thomas W.R.;
RT "Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I.";
RL Clin. Exp. Allergy 21:25-32 (1991).
RN [2]
RP SEQUENCE OF 98-309 FROM N.A.
RA Kent N., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 99-128.
RX MEDLINE=88229138; PubMed=3372999;
RA Lind P., Hansen O.C., Horn N.;
RT "The binding of mouse hybridoma and human IgE antibodies to the major fecal allergen, Derp I, of Dermatophagoides pteronyssinus. Relative binding site location and species specificity studied by solid-phase inhibition assays with radiolabeled antigen.";
RL J. Immunol. 140:4256-4262 (1988).
CC -!- FUNCTION: Thiol protease that hydrolyzes proteins, with a preference for phe or basic residues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of mite allergy are bronchial asthma, allergic rhinitis and conjunctivitis.
CC -!- SIMILARITY: Belongs to peptidase family C1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC

```

CC or send an email to license@isb-sib.ch).

CC EMBL; X65196; CAA46316.1; -  
 DR PIR; A27634; A27634.  
 DR HSP; P53634; IK3B.  
 DR MEROPS; C01.073; -  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; Pept\_cys\_acsite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPA1N.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SM00645; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE HIS; 1.  
 KW Allergen; Direct protein sequencing; Glycoprotein; Hydrolase; Signal;  
 KW Thiol protease; Zymogen.  
 FT SIGNAL 1 18 Potential.  
 FT PROPEP 19 98 Activation peptide.  
 FT CHAIN 99 321 Major mite fecal allergen Der f 1.  
 FT ACT\_SITE 133 133 By similarity.  
 FT ACT\_SITE 269 269 By similarity.  
 FT ACT\_SITE 288 288 By similarity.  
 FT CARBOHYD 151 151 N-linked (GLCNac...) (Potential).  
 FT DISULFID 102 216 By similarity.  
 FT DISULFID 130 170 By similarity.  
 FT DISULFID 164 202 By similarity.  
 FT CONFLICT 201 201 R -> Q (in Ref. 2).  
 FT CONFLICT 282 282 D -> V (in Ref. 2).  
 SQ SEQUENCE 321 AA; 36435 MW; 04523E54EBB476E CRC64;

Query Match 100.0%; Score 55; DB 1; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.023;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
 |||||  
 Db 265 QPNYHAVNIV 274

RESULT 6  
 BAC53948  
 ID BAC53948 PRELIMINARY; PRT; 321 AA.  
 AC BAC53948;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 Der f 1 allergen preproenzyme precursor.  
 OS Dermatophagoides farinae (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcotiformes; Astigmata; Analgoidea; Pyroglyphidae;  
 OC Dermatophagoides.  
 OX NCBI\_TaxID=6954;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yasuhara T., Takai T., Takahashi K., Yuuki T., Yokota T., Okudaira H.,  
 RA Okumura Y.;  
 RT "Dermatophagoides farinae Der f 1 allergen preproenzyme mRNA";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Dilworth R.J., Chuan K.Y., Thomas W.R.;  
 RT "Sequence analysis of cDNA coding for a major house dust mite  
 allergen, Der f 1,"  
 RL Clin. Exp. Allergy 21:25-32(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93357682; PubMed=8353459;  
 RA Chua K.Y., Kehal P.K., Thomas W.R.;  
 RT "Sequence polymorphisms of cDNA clones encoding the mite allergen Der  
 p 1,"  
 RL Int. Arch. Allergy Immunol. 101:364-368(1993).  
 DR EMBL; AB034946; BAC53948.1; -

KW Signal. 1 18 POTENTIAL.  
 FT SIGNAL 99 321  
 FT CHAIN 201 201  
 FT VARIANT 282 282 Q -> R (IN REF. 2).  
 FT VARIANT 282 282 V -> D (IN REF. 2).  
 SQ SEQUENCE 321 AA; 36391 MW; 83594754EBB4477 CRC64;  
 Query Match 100.0%; Score 55; DB 2; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 0.023;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10

|||||

Db 265 QPNYHAVNIV 274

RESULT 7

Q7M431  
 ID Q7M431 PRELIMINARY; PRT; 94 AA.  
 AC Q7M431;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Major fecal allergen Der p 1 (Fragments).  
 OS Dermatophagoides pteronyssinus (House-dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Dermatophagoides.  
 OX NCBI\_TaxID=6956;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=89098855; PubMed=2911558;  
 RA Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;  
 RT "Structural studies on the allergen Der p1 from the house dust mite  
 Dermatophagoides pteronyssinus: similarity with cysteine  
 proteinases,"  
 RL Protein Seq. Data Anal. 2:17-21(1989).  
 DR PIR; S03380; S03380.  
 DR GO; GO:0008234; F:cysteine-type peptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR Probom; PD000158; Peptidase\_C1; 1.  
 FT NON TER 1 1  
 FT NON TER 94 94  
 SQ SEQUENCE 94 AA; 10327 MW; 9BF744165C8428A8 CRC64;

Query Match 83.6%; Score 46; DB 2; Length 94;  
 Best Local Similarity 90.0%; Pred. No. 0.35;  
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
 |||||  
 Db 81 QPNYHAVNIV 90

RESULT 8

Q9FE41  
 ID Q9FE41 PRELIMINARY; PRT; 2876 AA.  
 AC Q9FE41;  
 DT 01-MAR-2001 (TrEMBLrel. 15, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)  
 DT 01-CCF-2004 (TrEMBLrel. 25, Last annotation update)  
 DE Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC  
 clone:PO433F09 (Oryza sativa (japonica cultivar-group) genomic DNA,  
 chromosome 1, PAC clone:PO441E11).  
 DE Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriatodeae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;



Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.

DR Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hijishita S., Honda M., Ichikawa Y., Idenuma A., Iijima M., Ikeda M.,  
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
RA Karasawa M., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,  
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun W.Y.,  
RA Yano M., Jiang J., Gojobori T.,  
RT "The genome sequence and structure of rice chromosome 1";  
RL Nature 420:312-316(2002).

DR EMBL; AF002539; BAB08213.2; -;  
DR EMBL; AF002521; BAA96774.2; -;  
DR Gramene; O9FE41; -;  
DR GO; GO:0003577; F:DNA binding; IEA.  
DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
DR GO; GO:0006310; P:DNA recombination; IEA.  
DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
DR InterPro; IPR005162; Retrotrans\_gag.  
DR InterPro; IPR002156; RNaseH.  
DR InterPro; IPR001594; Rve.  
DR InterPro; IPR000477; RVTse.  
DR Pfam; PF03732; Retrotrans\_gag; 1.  
DR Pfam; PF00075; RNaseH; 1.  
DR Pfam; PF00665; rve; 1.  
DR Pfam; PF00078; RVT; 2.  
KW RNA-directed DNA polymerase; Transferase.  
SQ SEQUENCE 2876 AA; 328575 MW; AEBBC34CB496764E CRC64;

Query Match 76.4%; Score 42; DB 2; Length 2876;  
Best Local Similarity 77.8%; Pred. No. 1e+02;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 PNYHAVNV 10  
DB 1057 PDYHAVNV 1065  
[1]|||||

RESULT 9  
Q8XX48 PRELIMINARY; PRT; 456 AA.  
AC Q8XX48;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE PROBABLE TRANSMEMBRANE PROTEIN.  
GN Name=RS01289; OrderedLocusNames=RS022270;  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GM11000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
RA Chandler M., Choine N., Claudel-Renard C., Cunac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
RA Signier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502(2002).

DR EMBL; AL646069; CAD15977.1; -;  
DR GO; GO:0016021; C:integral to membrane; IEA.  
KW Complete Proteome; Transmembrane.  
SQ SEQUENCE 456 AA; 50981 MW; 413DBE2154E5B79A CRC64;

Query Match 74.5%; Score 41; DB 2; Length 456;  
Best Local Similarity 77.8%; Pred. No. 20;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PNYHAVNV 10  
DB 102 PNYHAVSFV 110  
[1]|||||

RESULT 10  
RL31\_NEIMA  
ID RL31\_NEIMA STANDARD; PRT; 71 AA.  
AC Q9UR74;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE 50S ribosomal protein L31.  
GN Name=rpmE, OrderedLocusNames=NMA0495, NMB1956;  
OS Neisseria meningitidis (serogroup A), and  
OC Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
OC Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=65699, 491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;  
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;  
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,  
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,  
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,  
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
RA Whitehead S., Spratt B.G., Barrall B.G.;  
RT "Complete DNA sequence of a serogroup A strain of Neisseria  
RT meningitidis Z2491";  
RL Nature 404:502-506(2000).

RP SEQUENCE FROM N.A.  
RC STRAIN=MC58 / Serogroup B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,  
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,  
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,  
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,  
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,  
RA Ciftone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,  
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Masignani V., Pizzi M.,  
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,  
RA Venter J.C.;  
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain  
RT MC58";  
RL Science 287:1809-1815(2000).

CC -1- SIMILARITY: Belongs to the L31P family of ribosomal proteins.  
CC Subfamily A.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL162753; CAB83788.1; -;  
DR EMBL; AE002544; AAF42285.1; -;  
DR PIR; E81021; E81021.  
DR TIGR; NMB1956; -;  
DR HAVAP; MF\_00501; -; 1.

DR InterPro; IPR002150; Ribosomal\_L31.  
 DR Pfam; PF01197; Ribosomal\_L31; 1.  
 DR PRINTS; PRO1249; RIBOSOMAL\_L31.  
 DR TIGRFAMs; TIGR00105; L31; 1.  
 DR PROSITE; PS01143; RIBOSOMAL\_L31; 1.  
 DR Complete proteome; Ribosomal protein.  
 KW SEQUENCE 71 AA; 8118 MW; 59331C1E436792CC CRC64;

Query Match 72.7%; Score 40; DB 1; Length 71;  
 Best Local Similarity 75.0%; Pred. No. 4;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNYHAVNI 9  
 Db 7 PNYHEVNV 14

## RESULT 11

ACOH\_ERWHE STANDARD; PRT; 348 AA.  
 AC OS4459;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Phospho-2-dehydro-3-deoxyheptosate aldolase, Trp-sensitive  
 DE (EC 2.5.1.54) (Phospho-2-keto-3-deoxyheptosate aldolase) (DAHP  
 DE synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).  
 GN Name=aroh;  
 OS Erwinia herbicola.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Pantoea.  
 OX NCBI\_TaxID=549;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=98083064; PubMed=9422601;  
 RA Subramaniam P.S., Xie G., Xia T., Jensen R.A.;  
 RT "Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate  
 RT synthase from Neisseria gonorrhoeae in the context of its membership  
 RT in a protein family containing a subset of 3-deoxy-D-arabino-  
 RT heptulosonate 7-phosphate synthases.";  
 RL J. Bacteriol. 180:119-127(1998).  
 CC -1- FUNCTION: Stereospecific condensation of phosphoenolpyruvate (PEP)  
 CC and D-erythrose-4-phosphate (E4P) giving rise to 3-deoxy-D-  
 CC arabino-heptulosonate-7-phosphate (DAHP).  
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + D-erythrose 4-phosphate  
 CC + H(2)O = 3-deoxy-D-erythro-hept-2-ulosonate 7-phosphate +  
 CC phosphate.  
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC first step.

CC -1- SIMILARITY: Belongs to class-I DAHP synthetase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U93355; AAB96400.1; -.  
 CC HSPSP; P00886; IN8F.  
 DR InterPro; IPR006219; AroFGH.  
 DR InterPro; IPR006218; DAHP1/XDSA.  
 DR Pfam; PF00793; DAHP\_synth\_1; 1.  
 DR ProDom; PD005060; AroFGH; 1.  
 DR TIGRFAMs; TIGR00034; aroFGH; 1.  
 KW Aromatic amino acid biosynthesis; Transferase.  
 KW SEQUENCE 348 AA; 37744 MW; 1673534B92523B76 CRC64;

Query Match 70.9%; Score 39; DB 1; Length 348;  
 Best Local Similarity 66.7%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QPNYHAVNI 9  
 Db 237 QPNYHASDV 245

## RESULT 12

Q6XZ06 PRELIMINARY; PRT; 172 AA.  
 ID Q6XZ06;  
 AC Q6XZ06;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE YLBN.  
 OS Spiroplasma kunkelii.  
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;  
 OC Spiroplasmataceae; Spiroplasma.  
 OX NCBI\_TaxID=47834;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CR2-3x;  
 RC PubMed=15000753;  
 RA Zhao Y., Hammond R.W., Lee I.M., Roe B.A., Lin S., Davis R.E.;  
 RT "Cell division gene cluster in Spiroplasma kunkelii: functional  
 RT characterization offtsZ and the first report offtsA in mollicutes.";  
 RL DNA Cell Biol. 23:127-134(2004).  
 DR EMBL; AY198132; AAP42768.1; -.  
 SQ SEQUENCE 172 AA; 20015 MW; 90835BFABEAB9DEE CRC64;

Query Match 69.1%; Score 38; DB 2; Length 172;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNYHAVNI 10  
 Db 49 PNFNAINIV 57

## RESULT 13

AAP42768 PRELIMINARY; PRT; 172 AA.  
 ID AAP42768  
 AC AAP42768;  
 DT 23-APR-2004 (TREMBLrel. 27, Created)  
 DT 23-APR-2004 (TREMBLrel. 27, Last sequence update)  
 DT 23-APR-2004 (TREMBLrel. 27, Last annotation update)  
 DE YLBN.  
 OS Spiroplasma kunkelii.  
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;  
 OC Spiroplasmataceae; Spiroplasma.  
 OX NCBI\_TaxID=47834;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CR2-3x;  
 RC PubMed=15000753;  
 RA Zhao Y., Hammond R.W., Lee I.M., Roe B.A., Lin S., Davis R.E.;  
 RT "Cell division gene cluster in Spiroplasma kunkelii: functional  
 RT characterization offtsZ and the first report offtsA in mollicutes.";  
 RL DNA Cell Biol. 23:127-134(2004).  
 DR EMBL; AY198132; AAP42768.1; -.  
 SQ SEQUENCE 172 AA; 20015 MW; 90835BFABEAB9DEE CRC64;

Query Match 69.1%; Score 38; DB 2; Length 172;  
 Best Local Similarity 66.7%; Pred. No. 27;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNYHAVNI 10  
 Db 49 PNFNAINIV 57

## RESULT 14

YQFN\_BACSU STANDARD; PRT; 216 AA.  
 ID YQFN\_BACSU  
 AC P54471;

DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 01-OCT-2004 (Rel. 45, Last annotation update)  
DE Hypothetical protein yqfN.  
GN Name=yqfN; OrderedLocNames=BSU25130;  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RX MEDLINE=97124195; PubMed=8969508;  
RA Kizuno M., Masuda S., Takenaru K.-I., Hosono S., Sato T., Takeuchi M.,  
RA Kobayashi Y.;  
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
RT the Bacillus subtilis genome containing the skin element and many  
RT sporulation genes.";  
RL Microbiology 142:3103-3111(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entlan K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Goughly E.J., Grandi G.,  
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,  
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M.,  
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,  
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C.,  
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,  
RA Nobak M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
RA Park S.H., Parro V., Pohl T.M., Portetelie D., Porwolik S.,  
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,  
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,  
RA Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,  
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,  
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,  
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,  
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,  
RA Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis".  
RL Nature 390:249-256(1997).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D84432; BAA12491.1; -;  
DR EMBL; Z99116; CAB14448.1; -;  
DR PIR; H69953; H69953.  
DR Subtilisin; BG11660; yqfN.  
DR InterPro; IPR006901; DUF633.  
DR Pfam; PF04816; DUF633; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 216 AA; 23705 MW; 1FFB8A913D7A79D1 CRC64;

Query Match

69.1%; Score 38; DB 1; Length 216;

Best Local Similarity 77.8%; Pred. No. 35;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 QPNYHAVNI 9  
Db 99 QPNYHAVNI 107  
RESULT 15  
Q819P1 PRELIMINARY; PRT; 263 AA.  
AC Q819P1;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Derp1 antigen (fragment).  
OS Psoroptes ovis (Sheep scab mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Sarcoptoidea;  
OC Psoroptidae; Psoroptes.  
OX NCBI\_TaxID=83912;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22294898; PubMed=12406195;  
RA Lee A.J., Machell J., Van Den Broek A.H.M., Nisbet A.J.,  
RA Miller H.R.P., Isaac R.E., Huntley J.F.;  
RT "Identification of an antigen from the sheep scab mite, Psoroptes  
RT ovis, homologous with house dust mite group I allergens.";  
RL Parasite Immunol. 24:413-422(2002).  
DR EMBL; AF495854; AAC14671.1; -;  
DR MEROPS; C01.073; -;  
DR GO; GO:0004197; F: cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR ProDom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
FT NON TER 1 263  
FT NON TER 263 263  
SQ SEQUENCE 263 AA; 23576 MW; BF6DD21006DAB5B0 CRC64;  
Query Match 69.1%; Score 38; DB 2; Length 263;  
Best Local Similarity 75.0%; Pred. No. 43;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 3 NYHAVNIV 10  
Db 256 NFHAVNIV 263

Search completed: October 27, 2004, 17:50:18

Job time : 16.7809 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:26:32 ; Search time 13.7849 Seconds  
(without alignments)  
260.234 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	5 AAO20570	Aao20570 Cysteine
2	55	100.0	10	5 ABB98535	Abb98535 Cysteine
3	55	100.0	23	2 AAR34700	Aar34700 DFI-9(158
4	55	100.0	23	2 AAR36460	Aar36460 DFI-9(158
5	55	100.0	23	2 AAR51745	Aar51745 Der p I d
6	55	100.0	23	2 AAR51808	Aar51808 Der f I d
7	55	100.0	23	2 AAR71132	Aar71132 Dermatoph
8	55	100.0	23	2 AAW71925	Aaw71925 Dermatoph
9	55	100.0	23	2 AAW71983	Aaw71983 Dermatoph
10	55	100.0	23	2 AAV50437	Aav50437 Dermatoph
11	55	100.0	23	2 AAV50374	Aav50374 Dermatoph
12	55	100.0	23	4 AAU18977	Aau18977 T-cell ep
13	55	100.0	23	4 AAU19040	Aau19040 T-cell ep
14	55	100.0	25	2 AAR36474	Aar36474 DFI-27.1(
15	55	100.0	25	2 AAR36412	Aar36412 DFI-27.1(
16	55	100.0	25	2 AAR51822	Aar51822 Der f I d
17	55	100.0	25	2 AAW71996	Aaw71996 Dermatoph
18	55	100.0	25	2 AAV50451	Aav50451 Dermatoph
19	55	100.0	25	4 AAU19054	Aau19054 T-cell ep
20	55	100.0	72	4 AAU07749	Aau07749 House dus
21	55	100.0	211	2 AAY25677	Aay25677 Euroglyph
22	55	100.0	211	2 AAY25676	Aay25676 Euroglyph
23	55	100.0	211	2 AAY25678	Aay25678 Euroglyph
24	55	100.0	211	7 ADC34925	Adc34925 Euroglyph
25	55	100.0	211	7 ADC34926	Adc34926 Euroglyph

26	55	100.0	211	7 ADC34924	Adc34924 Euroglyph
27	55	100.0	212	2 AAY25679	Aay25679 Euroglyph
28	55	100.0	212	7 ADC34927	Adc34927 Euroglyph
29	55	100.0	222	2 AAR52742	Aar52742 Protein a
30	55	100.0	222	4 AAU07746	Aau07746 House dus
31	55	100.0	222	4 AAU07748	Aau07748 House dus
32	55	100.0	222	4 AAU07747	Aau07747 House dus
33	55	100.0	222	4 AAB98347	Aab98347 D. pteron
34	55	100.0	222	5 AABG67023	Abg67023 House dus
35	55	100.0	222	5 AAO20568	Aao20568 Cysteine
36	55	100.0	222	5 ABB98533	Abb98533 Cysteine
37	55	100.0	222	8 ADK52140	Adk52140 Der pl al
38	55	100.0	223	2 AAR78159	Aar78159 Recombina
39	55	100.0	223	4 AAB98328	Aab98328 D. farina
40	55	100.0	223	4 AAB98334	Aab98334 E. maynei
41	55	100.0	223	4 AAB98336	Aab98336 D. farina
42	55	100.0	223	4 AAB98345	Aab98345 D. pteron
43	55	100.0	223	4 AAB98331	Aab98331 D. pteron
44	55	100.0	223	4 AAB98339	Aab98339 D. farina
45	55	100.0	223	4 AAB98340	Aab98340 D. farina

#### ALIGNMENTS

##### RESULT 1

AAO20570  
ID AAO20570 standard; peptide; 10 AA.

XX AC AAO20570;

XX DT 02-JAN-2003 (first entry)

XX DE Cysteine protease epitope peptide region, SEQ ID No 4.

XX KW Antiallergic; antiinflammatory; antisthmatic; dermatological; allergen;  
anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
atopical eczema; epitope.

XX OS Dermatophagoides pteronyssinus.

XX PN WO200278736-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-FR001098.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PR 03-MAY-2001; 2001FR-00005929.

XX PR 29-MAY-2001; 2001US-00867159.

XX PA (ANTI-) ANTIALIS SARL.

XX PI Loria E, Terrasse G, Trehin Y;

XX WPI; 2002-750636/81.

XX Antiallergic compositions containing an anti-histamine, a histamine  
synthesis inhibitor, and optionally an allergen or nucleic acid coding  
for the allergen.

XX Claim 14; Page 11; 32pp; French.

XX The invention relates to antiallergic compositions containing an anti-  
histamine, a histamine synthesis inhibitor, and optionally an allergen or  
isolated nucleic acid molecule that has at least one polynucleotide  
sequence coding for the allergen, together with a pharmaceutical carrier.  
The pharmaceutical composition of the invention is useful as a non-  
specific antiallergic treatment, and also useful in the treatment of  
allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
allergic and atopical eczema. This sequence represents a peptide of a  
cysteine protease epitope region relating to the antiallergic

```

CC compositions of the invention
XX
SQ Sequence 10 AA;

Query Match      100.0%; Score 55; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 1 QPNYHAVNIV 10

RESULT 2
ABB98535
ID ABB98535 standard; peptide; 10 AA.
XX
AC ABB98535;
XX
XX 13-DEC-2002 (first entry)
XX
XX Cysteine protease epitope #2.
XX
XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunotherapy; allergen; allergic hypersensitivity reaction;
KW allergic asthma; allergic rhinitis; allergic atopic eczema;
KW cysteine protease.
XX
XX Dermatophagoides pteronyssinus.
XX
XX FR2822709-A1.
XX
XX 04-OCT-2002.
XX
XX 03-MAY-2001; 2001FR-00005929.
XX
XX 30-MAR-2001; 2001FR-00004370.
XX
XX (ANTI-) ANTIALIS SARR.
XX
XX Loria E, Terrasse G, Trehin Y;
XX
XX WPI; 2002-735037/80.
XX
XX Antiallergic composition, useful for preventing and treating e.g. asthma,
PT rhinitis or eczema, containing at least two of allergen, antihistamine
PT and histamine synthesis inhibitor.
XX
XX Claim 8; Page 6; 33pp; French.
XX
XX The present invention relates to an antiallergic pharmaceutical
CC composition (I) comprising a pharmaceutical carrier containing an active
CC agent combination of at least two of: an allergen; an antihistamine; and
CC a histamine synthesis inhibitor. (I) is used for treating or preventing
CC allergic hypersensitivity reactions, especially allergic asthma, allergic
CC rhinitis or allergic atopic eczema, in babies, children or adults. The
CC present sequence is a peptide fragment (epitope) of cysteine protease
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the
CC invention
XX
SQ Sequence 10 AA;

Query Match      100.0%; Score 55; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 1 QPNYHAVNIV 10

RESULT 3
AAR34700
ID AAR34700 standard; peptide; 23 AA.
XX
AC AAR34700;
XX
XX 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
XX DPI-9(158-180) a Dermatophagoides protein allergen.
XX
XX T cell epitope; house dust mite; allergy; soluble; Der pl.
XX
XX Synthetic.
XX
XX WO9308279-A1.
XX
XX 29-APR-1993.
XX
XX 15-OCT-1992; 92WO-US008637.
XX
XX 16-OCT-1991; 91US-00777859.
XX
XX 08-MAY-1992; 92US-00881396.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
XX WPI; 1993-152472/18.
XX
XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis
PT and treatment of sensitivity to house dust mite.
XX
XX Claim 10; Fig 3; 176pp; English.
XX
XX The peptide is one of a series of overlapping peptides synthesised by
CC standard techniques to cover the whole Dermatophagoides pteronyssinus Der
CC PI sequence. The T cell epitopes of the protein were mapped by detection
CC of the peptide's ability to stimulate T cell activity. The peptides may
CC be used for diagnosis and treatment of sensitivity to house dust mite
CC allergens. When administered to house dust mite sensitive individuals,
CC the peptides are capable of modifying the allergic response to the
CC allergens. The peptides may be modified for e.g. increasing solubility,
CC enhancing therapeutic or preventive efficacy or stability. See also
CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
XX Sequence 23 AA;

Query Match      100.0%; Score 55; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 9 QPNYHAVNIV 18

RESULT 4
AAR36460
ID AAR36460 standard; peptide; 23 AA.
XX
AC AAR36460;
XX
XX 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
XX DPI-9(158-180), a Dermatophagoides protein allergen.
XX
XX T cell epitope; house dust mite; allergy; soluble; Def pl.
XX
XX Synthetic.
XX
XX WO9308279-A1.
XX

```

PD 29-APR-1993.  
 XX  
 PF 15-OCT-1992; 92WO-US0008637.  
 XX  
 PR 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
 XX WPI; 1993-152472/18.  
 XX  
 PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 PT and treatment of sensitivity to house dust mite.  
 XX  
 PS Claim 44; Fig 4; 176pp; English.  
 CC  
 CC The peptide is one of a series of overlapping peptides synthesised by  
 CC standard techniques to cover the whole Dermatophagoides farinae Def PI  
 CC sequence. The T cell epitopes of the protein were mapped by detection of  
 CC the peptide's ability to stimulate T cell activity. The peptides may be  
 CC used for diagnosis and treatment of sensitivity to house dust mite  
 CC allergens. When administered to house dust mite sensitive individuals,  
 CC the peptides are capable of modifying the allergic response to the  
 CC allergens. The peptides may be modified for e.g. increasing solubility,  
 CC enhancing therapeutic or preventive efficacy or stability. See also  
 CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPNYHAVNIV 10  
 Db |||||  
 9 QPNYHAVNIV 18  
 RESULT 5  
 AAR51745  
 ID AAR51745 standard; protein; 23 AA.  
 AC AAR51745;  
 DT  
 DT 01-FEB-1995 (first entry)  
 XX  
 DE Der p I derived peptide, DP I-9(158-180).  
 XX  
 KW Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
 KW homology; D. farinae; Der f I; group II; Der p II; T-cell;  
 KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
 KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
 KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
 KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
 KW T cell subpopulations; unresponsive; immune response; tolerate.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 XX ZA9302677-A.  
 PN  
 XX 26-JAN-1994.  
 PD  
 PF 16-APR-1993; 93ZA-00002677.  
 XX  
 PR 16-APR-1993; 93ZA-00002677.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX WPI; 1993-152472/18.  
 XX  
 PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 PT and treatment of sensitivity to house dust mite.  
 XX  
 PS Claim 44; Fig 4; 176pp; English.  
 CC  
 CC The peptide is one of a series of overlapping peptides synthesised by  
 CC standard techniques to cover the whole Dermatophagoides farinae Def PI  
 CC sequence. The T cell epitopes of the protein were mapped by detection of  
 CC the peptide's ability to stimulate T cell activity. The peptides may be  
 CC used for diagnosis and treatment of sensitivity to house dust mite  
 CC allergens. When administered to house dust mite sensitive individuals,  
 CC the peptides are capable of modifying the allergic response to the  
 CC allergens. The peptides may be modified for e.g. increasing solubility,  
 CC enhancing therapeutic or preventive efficacy or stability. See also  
 CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPNYHAVNIV 10  
 Db |||||  
 9 QPNYHAVNIV 18  
 RESULT 6  
 AAR51808  
 ID AAR51808 standard; protein; 23 AA.  
 AC AAR51808;  
 DT  
 DT 01-FEB-1995 (first entry)  
 XX  
 DE Der f I derived peptide, DF I-9(158-180).  
 XX  
 KW Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
 KW homology; D. farinae; Der f I; group II; Der p II; T-cell;  
 KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
 KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
 KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
 KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
 KW T cell subpopulations; unresponsive; immune response; tolerate.  
 XX  
 OS Dermatophagoides farinae.  
 XX  
 XX ZA9302677-A.  
 PN  
 XX 26-JAN-1994.  
 PD  
 PF 16-APR-1993; 93ZA-00002677.  
 XX  
 PR 16-APR-1993; 93ZA-00002677.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX WPI; 1993-152472/18.  
 XX  
 PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 PT and treatment of sensitivity to house dust mite.  
 XX  
 PS Claim 44; Fig 4; 176pp; English.  
 CC  
 CC The sequences given in AAR51731-841 represent T-cell epitopes derived  
 CC from the group I and II protein allergens from the house dust mite D.  
 CC farinae and D. pteronyssinus. Der f I, Der f II, Der p I and Der p II  
 CC respectively. The Der f II proteinsCC shows high homology having an  
 CC identity of 88%, with an identity of 81% between the two group I proteins  
 CC (see also AAR51727-30). Fusion peptides may be produced which comprise at  
 CC least two or these antigenic fragments. Each region of these fusion  
 CC peptides may be derived from the same, or different, mite allergens. The  
 CC antigenic fragments may be altered by substitution, deletion or addition  
 CC to enhance their antigenicity. These peptides may be produced by chemical  
 CC synthesis, chemical cleavage of the protein allergen or by recombinant  
 CC techniques. These peptides, or the fusion peptides, when administered to  
 CC a house dust mite sensitive individual, are capable of modifying the  
 CC allergic response of the individual to the allergen. The peptides do not  
 CC bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the  
 CC full length protein allergen. This reduces the major complications of  
 CC standard immunotherapy, which are IgE-mediated responses such as  
 CC anaphylaxis. Exposure of mite allergic patients to these peptides may  
 CC tolerate or anergise appropriate T cell subpopulations such that they  
 CC become unresponsive to mite allergens and do not participate in mounting  
 CC an immune response upon exposure. Administration of the peptides may also  
 CC modify the lymphokine secretion profile as compared with exposure to the  
 CC naturally occurring mite protein allergen  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPNYHAVNIV 10  
 Db |||||  
 9 QPNYHAVNIV 18

PI Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX WPI; 1994-126807/15.  
 XX  
 PT Isolated and/or modified peptides comprising T-cell epitopes - of major  
 PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
 PT sensitivity to house dust mites.  
 XX  
 XX Disclosure; Fig 4; 154pp; English.  
 PS  
 CC The sequences given in AAR51731-841 represent T-cell epitopes derived  
 CC from the group I and II protein allergens from the house dust mite D.  
 CC farinae and D. pteronyssinus. Der f I, Der f II, Der p I and Der p II  
 CC respectively. The Der f II proteins show high homology having an  
 CC identity of 88%, with an identity of 81% between the two group I proteins  
 CC (see also AAR51727-30). Fusion peptides may be produced which comprise at  
 CC least two or these antigenic fragments. Each region of these fusion  
 CC peptides may be derived from the same, or different, mite allergens. The  
 CC antigenic fragments may be altered by substitution, deletion or addition  
 CC to enhance their antigenicity. These peptides may be produced by chemical  
 CC synthesis, chemical cleavage of the protein allergen or by recombinant  
 CC techniques. These peptides, or the fusion peptides, when administered to  
 CC a house dust mite sensitive individual, are capable of modifying the  
 CC allergic response of the individual to the allergen. The peptides do not  
 CC bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the  
 CC full length protein allergen. This reduces the major complications of  
 CC standard immunotherapy, which are IgE-mediated responses such as  
 CC anaphylaxis. Exposure of mite allergic patients to these peptides may  
 CC tolerate or anergise appropriate T cell subpopulations such that they  
 CC become unresponsive to mite allergens and do not participate in mounting  
 CC an immune response upon exposure. Administration of the peptides may also  
 CC modify the lymphokine secretion profile as compared with exposure to the  
 CC naturally occurring mite protein allergen  
 XX  
 XX Sequence 23 AA;  
 SQ  
 Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QPNYHAVNIV 10  
 Db |||||  
 9 QPNYHAVNIV 18  
 RESULT 7  
 AAR77132  
 ID AAR77132 standard; peptide; 23 AA.  
 XX  
 AC AAR77132;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 31-MAY-1996 (first entry)  
 XX  
 XX Dermatophagoides pteronyssinus group I peptide DPI-9 (158-180).  
 DE  
 XX House dust mite; DerpI; DerfI; DerpII; DerfII; allergen; allergy;  
 KW overlapping peptide; screening.  
 XX Dermatophagoides pteronyssinus.  
 OS  
 XX WO9528424-A1.  
 PN  
 XX 26-OCT-1995.  
 FD  
 XX 12-APR-1995; 95WO-US004481.  
 XX  
 XX 14-APR-1994; 94US-00227772.  
 PR  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Chen X, Evans S, Shaked Z, Franzen HM, Kuo M;  
 FI  
 XX

DR WPI; 1995-373765/48.  
 XX  
 PT Compsns. contg. house mite allergen-derived peptide(s), some of which are  
 PT new - are used to treat allergy, and are stable, soluble and able to  
 PT induce T cell non-responsiveness.  
 XX  
 XX Disclosure; Fig 2; 61pp; English.  
 PS  
 CC Claimed therapeutic compositions contain at least one of the peptides DPI  
 CC -21.2 and DPI-22.2 and also at least one of the new peptides DPI-23.31,  
 CC DPI-26.6, DPI-20.9, DPI-22.14 and DPI-25.15. The compositions are  
 CC useful for treating sensitivity to house dust mite allergens. The  
 CC peptides were identified by screening overlapping peptides derived from  
 CC D. pteronyssinus and D. farinae group I and II allergens for T-cell  
 CC reactivity in sensitised individuals. The present sequence is that of  
 CC overlapping peptide DPI-9 (158-180). (Updated on 25-MAR-2003 to correct  
 CC PR field.)  
 XX  
 XX Sequence 23 AA;  
 SQ  
 Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 QPNYHAVNIV 10  
 Db |||||  
 9 QPNYHAVNIV 18  
 RESULT 8  
 AAW71925  
 ID AAW71925 standard; peptide; 23 AA.  
 XX  
 AC AAW71925;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 16-DEC-1998 (first entry)  
 XX  
 XX Dermatophagoides Der p I protein peptide DPI-9.  
 DE  
 XX Genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
 KW Der p II; Der f I; Der f II; house dust mite allergy.  
 XX Dermatophagoides.  
 OS  
 XX US5820862-A.  
 PN  
 XX 13-OCT-1998.  
 PD  
 XX 07-JUN-1995; 95US-00482142.  
 PF  
 XX 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
 FI Shaked Z, Rogers BL;  
 XX  
 XX WPI; 1998-567590/48.  
 DR  
 XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.  
 PT  
 XX Disclosure; Col 89-90; 155pp; English.  
 PS  
 XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen selected  
 CC from the allergens Der p I, Der p II, Der f I, or Der f II. The invention



CC also describes modified peptides having similar or enhanced therapeutic  
 CC properties as the corresponding, naturally occurring allergen, but having  
 CC reduced side effects. AA71912 to AA72000, and AA72257 to AA72330  
 CC represent peptides from the present invention. (Updated on 25-MAR-2003 to  
 CC correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 23 AA;

Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0;

Qy 1 QPNYHAVNIV 10  
 Db 9 QPNYHAVNIV 18

RESULT 9  
 AA71983  
 ID AA71983 standard; peptide; 23 AA.

AC AA71983;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 16-DEC-1998 (first entry)

XX Dermatophagoides Der f I protein peptide DfI-9.

KW genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
 KW Der p II; Der f I; Der f II; house dust mite allergy.

XX Dermatophagoides.

OS US5820862-A.

PN 13-OCT-1998.

PD 07-JUN-1995; 95US-00482142.

PF 14-APR-1993; 93WO-US003471.

PR 14-APR-1994; 94US-00227772.

PR 19-MAY-1995; 95US-00445307.

XX (IMMU-) IMMULOGIC PHARM CORP.

PI Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;

PI Shaked Z, Rogers BL;

XX WPI; 1998-567590/48.

XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.

PS Disclosure; Col 121-122; 155pp; English.

XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen selected  
 CC from the allergens Der p I, Der p II, Der f I, or Der f II. The invention  
 CC also describes modified peptides having similar or enhanced therapeutic  
 CC properties as the corresponding, naturally occurring allergen, but having  
 CC reduced side effects. AA71912 to AA72000, and AA72257 to AA72330  
 CC represent peptides from the present invention. (Updated on 25-MAR-2003 to  
 CC correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 23 AA;

Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0;

Qy 1 QPNYHAVNIV 10  
 Db 9 QPNYHAVNIV 18

RESULT 10  
 AA50437

ID AA50437 standard; peptide; 23 AA.

AC AA50437;

DT 25-JAN-2000 (first entry)

XX Dermatophagoides sp major protein allergen Df I-9.

KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der f I; Der p I; Der p II; Der f II.

OS Dermatophagoides sp.

PN US5968526-A.

PD 19-OCT-1999.

PF 07-JUN-1995; 95US-00478572.

PR 14-APR-1994; 94US-00227772.

PR 12-APR-1995; 95WO-US004481.

PR 19-MAY-1995; 95US-00445307.

XX (IMMU-) IMMULOGIC PHARM CORP.

PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;

PI Chen X, Evans S, Kuo M;

XX WPI; 1999-590385/50.

XX Screening individuals for allergic reactions to T cell epitopes of major  
 PT allergens from house dust mites.

PS Disclosure; Col 121-122; 158pp; English.

XX This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with 1 or more  
 CC isolated T cell epitopes of the protein allergens I and II (DP I) and  
 CC (DP II) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the sensitivity  
 CC of the patient to house dust mites. (I) may be used to screen individuals  
 CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
 CC mite is a major cause of a variety of allergic disorders such as asthma,  
 CC rhinitis and ectopic dermatitis. AA50360-Y50542 and AA50546-Y50555  
 CC represent house dust mite allergen peptide fragments derived from Der p  
 CC I, Der f II, Der f I and Der f II

XX Sequence 23 AA;

Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0;

Qy 1 QPNYHAVNIV 10  
 Db 9 QPNYHAVNIV 18

RESULT 11  
 AA50374

ID AAY50374 standard; peptide; 23 AA.  
 AC AAY50374;  
 XX  
 DT 25-JAN-2000 (first entry)  
 XX  
 XX Dermatophagoides sp major protein allergen DP I-9.  
 XX  
 XX Allergen; house dust mite; detection; sensitivity; T cell epitope;  
 KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
 KW Der f I; Der p I; Der p II; Der f II.  
 XX  
 OS Dermatophagoides sp.  
 XX  
 XX US5968526-A.  
 FN  
 XX 19-OCT-1999.  
 PD  
 XX 07-JUN-1995; 95US-00478572.  
 XX  
 XX 14-APR-1994; 94US-00227772.  
 PR  
 XX 12-APR-1995; 95WO-US004481.  
 PR  
 XX 19-MAY-1995; 95US-00445307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 XX Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;  
 PI Chen X, Evans S, Kuo M;  
 XX  
 XX WPI; 1999-590385/50.  
 DR  
 XX Screening individuals for allergic reactions to T cell epitopes of major  
 PT allergens from house dust mites.  
 FT  
 XX Claim In; Col 91-92; 158pp; English.  
 PS  
 XX This invention describes a novel method (I) for detecting whether an  
 CC individual is sensitive to Dermatophagoides (house dust mites). The  
 CC method involves detecting sensitivity to house dust mites in patients,  
 CC comprising combining a blood sample from the individual with I or more  
 CC isolated T cell epitopes of the protein allergens I and II (DP I) and  
 CC (DP II) from Dermatophagoides (house dust mites). 32 T cell epitopes  
 CC with varying, defined amino acids sequences (given in the specification)  
 CC may be used in (I). The sample and allergens are combined under  
 CC conditions appropriate for the binding of blood components with the  
 CC polypeptides. The extent of binding is then indicative of the sensitivity  
 CC of the patient to house dust mites. (I) may be used to screen individuals  
 CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
 CC mite is a major cause of a variety of allergic disorders such as asthma,  
 CC rhinitis and ectopic dermatitis. AAY50360-Y50542 and AAY50546-Y50555  
 CC represent house dust mite allergen peptide fragments derived from Der p  
 CC I, Der f II, Der p I and Der f II  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPNYHAVNIV 10  
 DB 9 QPNYHAVNIV 18  
 RESULT 12  
 AAU18977  
 ID AAU18977 standard; peptide; 23 AA.  
 XX  
 AC AAU18977;  
 XX  
 XX 04-DEC-2001 (first entry)  
 DT  
 XX T-cell epitope containing peptide DPI-9.  
 DE

XX House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 XX US6268491-B1.  
 FN  
 XX 31-JUL-2001.  
 PD  
 XX 07-JUN-1995; 95US-00484296.  
 XX  
 XX 16-OCT-1991; 91US-00777859.  
 PR  
 XX 08-MAY-1992; 92US-00881396.  
 PR  
 XX 14-APR-1993; 93WO-US003471.  
 PR  
 XX 14-APR-1994; 94US-00227772.  
 PR  
 XX 19-MAY-1995; 95US-00445307.  
 XX  
 PA (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 XX  
 XX WPI; 2001-549074/61.  
 DR  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX  
 XX Example 3; Fig 3; 158pp; English.  
 PS  
 XX The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic T-cell epitope containing peptide  
 CC derived from the Dermatophagoides allergenic proteins  
 XX  
 SQ Sequence 23 AA;  
 Query Match 100.0%; Score 55; DB 4; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00088;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPNYHAVNIV 10  
 DB 9 QPNYHAVNIV 18  
 RESULT 13  
 AAU19040  
 ID AAU19040 standard; peptide; 23 AA.  
 XX  
 XX AAU19040;  
 AC  
 XX 04-DEC-2001 (first entry)  
 DT  
 XX T-cell epitope containing peptide DFI-9.  
 DE  
 XX House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX

OS Dermatophagoides farinae.  
XX US6268491-B1.  
XX 31-JUL-2001.  
XX 07-JUN-1995; 95US-00484296.  
XX 16-OCT-1991; 91US-00777859.  
XX 08-MAY-1992; 92US-00881396.  
XX 14-APR-1993; 93WO-US003471.  
XX 14-APR-1994; 94US-00227772.  
XX 19-MAY-1995; 95US-00445307.  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
PI Evans S, Shaked Z;  
PI WPI; 2001-549074/61.  
XX Peptides comprising T cell groups of the major allergens from  
PT Dermatophagoides (house dust mites), useful for treating house dust mite  
PT allergy in humans, and for diagnosing sensitivity to house dust mite  
PT protein allergens.  
XX Claim 5; Fig 4; 158pp; English.  
XX The invention relates to an isolated peptide of the major protein  
CC allergens of the genus Dermatophagoides, which comprises at least one T  
CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
CC or DF II. The isolated peptide comprises at least two regions, each  
CC region comprising at least one T cell group of a protein allergen of the  
CC genus Dermatophagoides. The regions are derived from the same or  
CC different protein allergens of the genus Dermatophagoides. The peptides  
CC are useful for treating house dust mite allergy in humans. The peptides  
CC are also useful for detecting or diagnosing sensitivity to house dust  
CC mite protein allergens. The present peptides have similar or enhanced  
CC therapeutic properties as the naturally-occurring allergen, but have  
CC reduced side effects, and increased solubility and stability. The present  
CC sequence represents an allergenic T-cell epitope containing peptide  
CC derived from the Dermatophagoides allergenic proteins  
XX  
XX SQ Sequence 23 AA;  
Query Match 100.0%; Score 55; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.00088;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 QPNYHAVNIV 10  
DB 9 QPNYHAVNIV 18  
RESULT 14  
AAR36474  
ID AAR36474 standard; peptide; 25 AA.  
XX  
XX AAR36474;  
XX 25-MAR-2003 (revised)  
DT 12-AUG-1993 (first entry)  
XX  
XX DFI-27.1(161-185), a Dermatophagoides protein allergen.  
DE  
XX  
XX T cell epitope; house dust mite; allergy; soluble; Def pI.  
XX  
XX Synthetic.  
XX WO9308279-A1.  
XX  
XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
XX and treatment of sensitivity to house dust mite.  
XX Claim 10; Fig 3; 176pp; English.

PF 15-OCT-1992; 92WO-US008637.  
XX  
XX 16-OCT-1991; 91US-00777859.  
XX 08-MAY-1992; 92US-00881396.  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
PI WPI; 1993-152472/18.  
XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
XX and treatment of sensitivity to house dust mite.  
XX Claim 4; Fig 4; 176pp; English.  
XX The peptide is one of a series of overlapping peptides synthesised by  
CC standard techniques to cover the whole Dermatophagoides farinae Def pI  
CC sequence. The T cell epitopes of the protein were mapped by detection of  
CC the peptide's ability to stimulate T cell activity. The peptides may be  
CC used for diagnosis and treatment of sensitivity to house dust mite  
CC allergens. When administered to house dust mite sensitive individuals,  
CC the peptides are capable of modifying the allergic response to the  
CC allergens. The peptides may be modified for e.g. increasing solubility,  
CC enhancing therapeutic or preventive efficacy or stability. See also  
CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX  
XX SQ Sequence 25 AA;  
Query Match 100.0%; Score 55; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.00097;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 QPNYHAVNIV 10  
DB 6 QPNYHAVNIV 15  
RESULT 15  
AAR36412  
ID AAR36412 standard; peptide; 25 AA.  
XX  
XX AAR36412;  
XX  
XX 25-MAR-2003 (revised)  
DT 12-AUG-1993 (first entry)  
XX  
XX DFI-27.1(161-185) a Dermatophagoides protein allergen.  
DE  
XX  
XX T cell epitope; house dust mite; allergy; soluble; Der pI.  
XX  
XX Synthetic.  
XX WO9308279-A1.  
XX  
XX 29-APR-1993.  
XX  
XX 15-OCT-1992; 92WO-US008637.  
XX  
XX 16-OCT-1991; 91US-00777859.  
XX 08-MAY-1992; 92US-00881396.  
XX (IMMU-) IMMULOGIC PHARM CORP.  
XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
PI WPI; 1993-152472/18.  
XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
XX and treatment of sensitivity to house dust mite.  
XX Claim 10; Fig 3; 176pp; English.

XX The peptide is one of a series of overlapping peptides synthesised by  
 CC standard techniques to cover the whole Dermatophagoides pteronyssinus Der  
 CC PI sequence. The T cell epitopes of the protein were mapped by detection  
 CC of the peptide's ability to stimulate T cell activity. The peptides may  
 CC be used for diagnosis and treatment of sensitivity to house dust mite  
 CC allergens. When administered to house dust mite sensitive individuals,  
 CC the peptides are capable of modifying the allergic response to the  
 CC allergens. The peptides may be modified for e.g. increasing solubility,  
 CC enhancing therapeutic or preventive efficacy or stability. See also  
 CC AAE34686-700 and AAE36398-490. (Updated on 25-MAR-2003 to correct PN  
 CC field.)

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 55; DB 2; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 0.00097;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10  
 Db 6 QPNYHAVNIV 15

Search completed: October 27, 2004, 17:43:57  
 Job time : 13.7849 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:50:34 ; Search time 10.3586 Seconds  
(without alignments)  
312.993 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	10	US-09-867-159A-4
2	55	100.0	210	14	US-10-001-245-185
3	55	100.0	211	10	US-09-847-208-95
4	55	100.0	211	14	US-10-001-245-184
5	55	100.0	222	10	US-09-867-159A-2
6	55	100.0	222	14	US-10-001-245-88
7	55	100.0	246	14	US-10-001-245-181
8	55	100.0	320	9	US-09-877-160-2
9	55	100.0	320	10	US-09-847-208-79
10	55	100.0	320	14	US-10-001-245-179
11	55	100.0	321	10	US-09-847-208-73
12	55	100.0	321	14	US-10-001-245-180
13	55	100.0	321	14	US-10-001-245-183

```

14 55 100.0 327 14 US-10-001-245-182 Sequence 182, App
15 47 85.5 222 14 US-10-001-245-14 Sequence 14, Appl
16 47 85.5 222 14 US-10-001-245-16 Sequence 16, Appl
17 47 85.5 222 14 US-10-001-245-18 Sequence 18, Appl
18 47 85.5 222 14 US-10-001-245-20 Sequence 20, Appl
19 47 85.5 222 14 US-10-001-245-22 Sequence 22, Appl
20 47 85.5 222 14 US-10-001-245-24 Sequence 24, Appl
21 47 85.5 222 14 US-10-001-245-26 Sequence 26, Appl
22 47 85.5 222 14 US-10-001-245-28 Sequence 28, Appl
23 47 85.5 222 14 US-10-001-245-30 Sequence 30, Appl
24 47 85.5 222 14 US-10-001-245-32 Sequence 32, Appl
25 47 85.5 222 14 US-10-001-245-34 Sequence 34, Appl
26 42 76.4 723 16 US-10-437-963-193170 Sequence 193170,
27 42 76.4 2683 16 US-10-437-963-185370 Sequence 185370,
28 42 76.4 2771 16 US-10-437-963-195406 Sequence 195406,
29 42 76.4 2882 16 US-10-437-963-195412 Sequence 195412,
30 39 70.9 76 15 US-10-424-599-240335 Sequence 240335,
31 38 69.1 461 14 US-10-341-200-31 Sequence 31, Appl
32 37 67.3 49 16 US-10-437-963-155821 Sequence 155821,
33 37 67.3 77 15 US-10-424-599-147667 Sequence 147667,
34 37 67.3 101 15 US-10-424-599-236670 Sequence 236670,
35 37 67.3 176 15 US-10-424-599-193859 Sequence 193859,
36 37 67.3 236 14 US-10-029-386-33438 Sequence 33438, A
37 37 67.3 901 16 US-10-408-765A-1259 Sequence 1259, Ap
38 37 67.3 903 16 US-10-408-765A-1332 Sequence 1332, Ap
39 37 67.3 1652 14 US-10-369-493-3223 Sequence 3223, Ap
40 36 65.5 130 16 US-10-437-963-109404 Sequence 109404,
41 36 65.5 144 15 US-10-424-599-148912 Sequence 148912,
42 36 65.5 155 15 US-10-425-114-66667 Sequence 66667, A
43 36 65.5 150 16 US-10-437-963-187625 Sequence 187625,
44 36 65.5 205 15 US-10-425-114-66321 Sequence 66321, A
45 36 65.5 207 15 US-10-425-114-64210 Sequence 64210, A

```

#### ALIGNMENTS

##### RESULT 1

US-09-867-159A-4

; Sequence 4, Application US/09867159A

; Publication No. US20030104013A1

; GENERAL INFORMATION:

; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE

; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one

; FILE REFERENCE: B112812US-antialis TREHIN, YVES

; CURRENT APPLICATION NUMBER: US/09/867,159A

; PRIOR FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: FR01/04370

; PRIOR FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: FR01/05929

; PRIOR FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Dermatophagoides pteronyssinus

; FEATURE:

; NAME/KEY: peptide

; LOCATION: (1)..(10)

; OTHER INFORMATION: Comprises epitope from cystine protease.

US-09-867-159A-4

Query Match 100.0%; Score 55; DB 10; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0014;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10

Db 1 QPNYHAVNIV 10

```

RESULT 2
US-10-001-245-185
; Sequence 185, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 185
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-001-245-185

Query Match      100.0%; Score 55; DB 14; Length 210;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10
| | | | | | | |
DB 166 QPNYHAVNIV 175

RESULT 3
US-09-847-208-95
; Sequence 95, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Euroglyphus maynei (House-dust mite)
US-09-847-208-95

Query Match      100.0%; Score 55; DB 10; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10
| | | | | | | |
DB 167 QPNYHAVNIV 176

RESULT 4
US-10-001-245-184
; Sequence 184, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 185
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Euroglyphus maynei
US-10-001-245-184

Query Match      100.0%; Score 55; DB 10; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10
| | | | | | | |
DB 166 QPNYHAVNIV 175

RESULT 5
US-09-867-159A-2
; Sequence 2, Application US/09867159A
; Publication No. US20030104013A1
; GENERAL INFORMATION:
; APPLICANT: ANTIALIS
; APPLICANT: TERRASSE, GAETAN
; APPLICANT: LORIA, EMILE
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one
; TITLE OF INVENTION: and at least one anti-histamine compound
; FILE REFERENCE: B112812US-antialis
; CURRENT APPLICATION NUMBER: US/09/867,159A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: FR01/04370
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: FR01/05929
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(222)
; OTHER INFORMATION: peptide sequence from cystine protease.
US-09-867-159A-2

Query Match      100.0%; Score 55; DB 10; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10
| | | | | | | |
DB 166 QPNYHAVNIV 175

RESULT 6
US-10-001-245-88
; Sequence 88, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245

```

```

; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 88
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-88

Query Match      100.0%; Score 55; DB 14; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNYHAVNIV 10
DB      166 QPNYHAVNIV 175

RESULT 7
US-10-001-245-181
; Sequence 181, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 181
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Euroglyphus maynei
US-10-001-245-181

Query Match      100.0%; Score 55; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNYHAVNIV 10
DB      190 QPNYHAVNIV 199

RESULT 8
US-09-877-160-2
; Sequence 2, Application US/09877160
; Publication No. US20020197268A1
; GENERAL INFORMATION:
; APPLICANT: Ching-Hsaiing, Hsu
; APPLICANT: Cheng, Winston T. K.
; TITLE OF INVENTION: ALLERGEN-CONTAINING MILK FOR ALLERGY
; FILE REFERENCE: 12774-003001
; CURRENT APPLICATION NUMBER: US/09/877,160
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 320
; TYPE: PRT

```

```

; ORGANISM: Dermatophagoides pteronyssinus
US-09-877-160-2

Query Match      100.0%; Score 55; DB 9; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNYHAVNIV 10
DB      264 QPNYHAVNIV 273

RESULT 9
US-09-847-208-79
; Sequence 79, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67,002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus (House-dust mite)
US-09-847-208-79

Query Match      100.0%; Score 55; DB 10; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNYHAVNIV 10
DB      264 QPNYHAVNIV 273

RESULT 10
US-10-001-245-179
; Sequence 179, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 179
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-179

Query Match      100.0%; Score 55; DB 14; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QPNYHAVNIV 10

```

```

Db      264 QPNYHAVNIV 273

RESULT 11
US-09-847-208-73
; Sequence 73, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae (House-dust mite)
US-09-847-208-73

Query Match      100.0%; Score 55; DB 10; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QPNYHAVNIV 10
      |||||
Db      265 QPNYHAVNIV 274

RESULT 12
US-10-001-245-180
; Sequence 180, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 180
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Euroglyphus maynei
US-10-001-245-180

Query Match      100.0%; Score 55; DB 14; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QPNYHAVNIV 10
      |||||
Db      265 QPNYHAVNIV 274

RESULT 13
US-10-001-245-183
; Sequence 183, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 180
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Euroglyphus maynei
US-10-001-245-180

Query Match      100.0%; Score 55; DB 14; Length 321;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QPNYHAVNIV 10
      |||||
Db      265 QPNYHAVNIV 274

RESULT 14
US-10-001-245-182
; Sequence 182, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Euroglyphus maynei
US-10-001-245-182

Query Match      100.0%; Score 55; DB 14; Length 327;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QPNYHAVNIV 10
      |||||
Db      271 QPNYHAVNIV 280

RESULT 15
US-10-001-245-14
; Sequence 14, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2

```



```
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-14

Query Match      85.5%; Score 47; DB 14; Length 222;
Best Local Similarity 90.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 QPNYHAVNIV 10
Db      166 QPNYHAVNIV 175

Search completed: October 27, 2004, 18:15:44
Job time : 11.3586 secs
```





Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00077;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
 Db 9 QPNYHAVNIV 18

## RESULT 2

US-08-482-142-86  
 ; Sequence 86, Application US/08482142  
 ; Patent No. 5820862  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/482,142  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/445,307  
 ; FILING DATE: 07 June 1995

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 86:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-482-142-86

Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00077;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
 Db 9 QPNYHAVNIV 18

## RESULT 3

US-08-478-572-23  
 ; Sequence 23, Application US/08478572

; Patent No. 5968526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207

; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/478,572  
 ; FILING DATE: 07-June-1995

; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 23:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-478-572-23

Query Match 100.0%; Score 55; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.00077;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
 Db 9 QPNYHAVNIV 18

## RESULT 4

US-08-478-572-86  
 ; Sequence 86, Application US/08478572  
 ; Patent No. 5968526

; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev

; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE: 07-June-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-478-572-86

Query Match 100.0%; Score 55; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 9 QPNYHAVNIV 18

RESULT 5
US-08-484-296-23
; Sequence 23, Application US/08484296
; Patent No. 6268491
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,572
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-484-296-23

Query Match 100.0%; Score 55; DB 3; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 9 QPNYHAVNIV 18

RESULT 6
US-08-484-296-86
; Sequence 86, Application US/08484296
; Patent No. 6268491
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,296
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; CURRENT APPLICATION DATA:
```

INFORMATION FOR SEQ ID NO: 86;  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-484-296-86

Query Match 100.0%; Score 55; DB 3; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.00077;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
Db 9 QPNYHAVNIV 18

RESULT 7  
PCT-US95-04481-14  
Sequence 14, Application PC/TUS9504481  
GENERAL INFORMATION:

APPLICANT: Garman, Richard  
TITLE OF INVENTION: Pharmaceutical Peptide Formulations For Treatment of Dust Mit  
NUMBER OF SEQUENCES: 54  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04481  
FILING DATE:  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/227,772  
FILING DATE: April 14, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Vanstone, Darlene A.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 017.5 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
PCT-US95-04481-14

Query Match 100.0%; Score 55; DB 5; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.00077;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
Db 9 QPNYHAVNIV 18

RESULT 8  
US-08-482-142-100  
Sequence 100, Application US/08482142  
Patent No. 5820862  
GENERAL INFORMATION:

APPLICANT: Garman, Richard  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154

APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,142  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.

REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-482-142-100

Query Match 100.0%; Score 55; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.00085;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10  
Db 6 QPNYHAVNIV 15

RESULT 9  
US-08-478-572-100  
Sequence 100, Application US/08478572  
Patent No. 5968526  
GENERAL INFORMATION:

APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-478-572-100

Query Match 100.0%; Score 55; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.00085;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10  
DB 6 QPNYHAVNIV 15

RESULT 10  
US-08-484-296-100  
Sequence 100, Application US/08484296  
Patent No. 6268491  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,296  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE: 07 June 1995  
ATTORNEY/AGENT INFORMATION:

NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017.6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-484-296-100

Query Match 100.0%; Score 55; DB 3; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.00085;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNYHAVNIV 10  
DB 6 QPNYHAVNIV 15

RESULT 11  
US-07-945-288-11  
Sequence 11, Application US/97945288  
Patent No. 5433948  
GENERAL INFORMATION:  
APPLICANT: Thomas, Wayne R.  
APPLICANT: Chua, Kaw-Yan  
TITLE OF INVENTION: CLONING AND SEQUENCING OF ALLERGENS FROM  
DERMATOPHAGOIDES (HOUSE DUST MITES)  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/945,288  
FILING DATE: 19920910  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 936,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 50  
OTHER INFORMATION: /label-Xaa is His or Tyr  
FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: 81
; OTHER INFORMATION: /label=Xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label=Xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label=Xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label=Xaa is Glu or Gln
; US-07-945-288-11

```

```

Query Match      100.0%; Score 55; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. NO. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QPNYHAVNIV 10
Db      166 QPNYHAVNIV 175

```

```

RESULT 12
US-08-462-831-11
; Sequence 11, Application US/08462831
; Patent No. 5552142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,831
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature

```

```

; LOCATION: 50
; OTHER INFORMATION: /label=Xaa is His or Tyr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 81
; OTHER INFORMATION: /label=Xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label=Xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label=Xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label=Xaa is Glu or Gln
; US-08-462-831-11

```

```

Query Match      100.0%; Score 55; DB 1; Length 222;
Best Local Similarity 100.0%; Pred. NO. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 QPNYHAVNIV 10
Db      166 QPNYHAVNIV 175

```

```

RESULT 13
US-08-461-809-11
; Sequence 11, Application US/08461809
; Patent No. 5770202
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM
; DERMATOPHAGOIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,809
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/945,288
; FILING DATE: 10 SEPTEMBER 1992
; APPLICATION NUMBER: US 580,655
; FILING DATE: 11 SEPTEMBER 1990
; APPLICATION NUMBER: US 458,642
; FILING DATE: 13 FEBRUARY 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MANDRAGOURAS, AMY E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```



MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 50  
OTHER INFORMATION: /label=Xaa is His or Tyr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 81  
OTHER INFORMATION: /label=Xaa is Glu or Lys  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 124  
OTHER INFORMATION: /label=Xaa is Ala or Val  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 136  
OTHER INFORMATION: /label=Xaa is Ser or Thr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 215  
OTHER INFORMATION: /label=Xaa is Glu or Gln  
US-08-461-809-11  
Query Match 100.0%; Score 55; DB 1; Length 222;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QPNYHAVNIV 10  
Db 166 QPNYHAVNIV 175  
RESULT 14  
US-08-461-441-11  
Sequence 11, Application US/08461441  
Patent No. 5773002  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,441  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/945,288  
FILING DATE: 10 SEPTEMBER 1992  
APPLICATION NUMBER: US 580,655  
FILING DATE: 11 SEPTEMBER 1990  
APPLICATION NUMBER: US 458,642  
FILING DATE: 13 FEBRUARY 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid

LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 50  
OTHER INFORMATION: /label=Xaa is His or Tyr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 81  
OTHER INFORMATION: /label=Xaa is Glu or Lys  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 124  
OTHER INFORMATION: /label=Xaa is Ala or Val  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 136  
OTHER INFORMATION: /label=Xaa is Ser or Thr  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 215  
OTHER INFORMATION: /label=Xaa is Glu or Gln  
US-08-461-441-11  
Query Match 100.0%; Score 55; DB 1; Length 222;  
Best Local Similarity 100.0%; Pred. No. 0.01;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QPNYHAVNIV 10  
Db 166 QPNYHAVNIV 175  
RESULT 15  
PCT-US93-08518-11  
Sequence 11, Application PC/TUS9308518  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS FROM  
TITLE OF INVENTION: DERMATOPHAGOIDES  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08518  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/945,288  
FILING DATE: 10 SEPTEMBER 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: MANDRAGOURAS, AMY E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IPC-010CC (IMI-024)  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 50
; OTHER INFORMATION: /label-xaa is His or Tyr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 81
; OTHER INFORMATION: /label-xaa is Glu or Lys
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 124
; OTHER INFORMATION: /label-xaa is Ala or Val
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 136
; OTHER INFORMATION: /label-xaa is Ser or Thr
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 215
; OTHER INFORMATION: /label-xaa is Glu or Gln
; PCT-US93-08518-11
```

```

Query Match      100.0%; Score 55; DB 5; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      1 QPNYHAVNIV 10
Db      166 QPNYHAVNIV 175
```

```

Search completed: October 27, 2004, 17:53:01
Job time : 3.54582 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:11:24 ; Search time 22.5 Seconds  
(without alignments)  
42.763 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55  
Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	32.7	10	2	cytochrome-c oxida
2	17	30.9	8	2	leucokinin VI - Ma
3	16	29.1	7	1	hypothalamic hepta
4	16	29.1	9	2	Ig heavy chain CRD
5	16	29.1	9	2	phosphoenolpyruvat
6	16	29.1	10	2	6-phosphofructo-2-
7	16	29.1	10	2	alpha-2-macroglobu
8	16	29.1	10	2	Ig heavy chain CRD
9	16	29.1	10	2	peptide-N4-(N-acet
10	16	29.1	10	2	neurokinin A - rat
11	16	29.1	10	2	neurokinin A - Acl
12	15	27.3	7	2	amine oxidase (cop
13	15	27.3	7	2	choline oxidase (B
14	15	27.3	8	2	A44960
15	15	27.3	8	2	neuropeptide Led-C
16	15	27.3	8	2	hypertrehalosemic
17	15	27.3	8	2	hypertrehalosemic
18	15	27.3	8	2	adipokinetic hormo
19	15	27.3	8	2	adipokinetic hormo
20	15	27.3	8	2	adipokinetic hormo
21	15	27.3	8	2	A43776
22	15	27.3	8	2	hypertrehalosemic
23	15	27.3	8	2	hypertrehalosemic
24	15	27.3	8	2	adipokinetic hormo
25	15	27.3	9	2	neuropeptide M-I -
26	15	27.3	9	2	amine oxidase (cop
27	15	27.3	9	2	Ig heavy chain CRD
28	15	27.3	9	2	phosphoenolpyruvat
29	15	27.3	10	1	cytochrome-c oxida
					tachykinin IV - mi

30 15 27.3 10 2 JCL1416  
31 15 27.3 10 2 S09138  
32 15 27.3 10 2 PT0213  
33 15 27.3 10 2 JQ0943  
34 14 25.5 6 2 PQ0008  
35 14 25.5 7 2 S71857  
36 14 25.5 8 2 A59028  
37 14 25.5 9 2 A61230  
38 14 25.5 9 2 C41170  
39 14 25.5 9 2 PT0288  
40 14 25.5 9 2 PT0324  
41 14 25.5 9 2 D41978  
42 14 25.5 10 2 S10926  
43 14 25.5 10 2 A60647  
44 13 23.6 6 2 A44916  
45 13 23.6 7 2 S21230

hypertrehalosemic  
hypertrehalosemic  
T-cell receptor al  
hypothetical 1.3K  
angiotensin-conver  
glutathione transf  
MHC class I histoc  
calsequestrin, car  
photosystem II pro  
Ig heavy chain CRD  
Ig heavy chain CRD  
callipyrinamide 4 -  
inhibin beta-A cha  
neurotrophin C - bov  
mosquitocidal toxi  
dermorphin (Trp-4,

## ALIGNMENTS

### RESULT 1

S65387  
cytochrome-c oxidase (EC 1.9.3.1) chain VII b, cardiac - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 09-Jul-2004  
C:Accession: S65387; S65386  
R:Schaeffer, H.; Noack, H.; Hallang, W.; Brandt, U.; von Jagow, G.  
Eur. J. Biochem. 230, 235-241, 1995  
A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-termi  
A:Reference number: S65372; MUID:95324529; PMID:7601105  
A:Accession: S65387  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SCH>  
A:Cross-references: UNIPROT:P80431  
A:Accession: S65386  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <SC>  
C:Keywords: cardiac muscle; heart; oxidoreductase

Query Match 32.7%; Score 18; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYH 5  
DB 7 PTFH 10

### RESULT 2

JS0316  
leucokinin VI - Madeira cockroach  
C:Species: Leucophaea maderae (Madeira cockroach)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
R:Holman, G.M.; Cock, B.J.; Nachman, R.J.  
Comp. Biochem. Physiol. C 88, 27-30, 1987  
A:Title: Isolation, primary structure, and synthesis of leucokinin V and VI: myotropic i  
A:Reference number: JS0315  
A:Accession: JS0316  
A:Molecule type: protein  
A:Residues: 1-8 <HOL>  
A:Cross-references: UNIPROT:P19988  
C:Comment: Leucokinin, a family of cephalomyotropic peptides, stimulate contractile acti  
C:Keywords: amidated carboxyl end; cephalomyotropic peptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 30.9%; Score 17; DB 2; Length 8;  
Best Local Similarity 33.3%; Pred. No. 2.8e+03;  
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

Query Match      29.1%; Score 16; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 QPNYH 5
      |||
Db      1 QPIIH 5

RESULT 6
A:Accession: A43405
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: A43405
R:Ventura, F.; Rosa, J.L.; Ambrosio, S.; Pilakis, S.J.; Bartrons, R.
A:Title: Isolation, structure and synthesis of a heptapeptide with in vitro ACTH-releasing activity
A:Reference number: A01417; PMID:81213980; PMID:6263778
A:Molecule type: protein
A:Residues: 1-7 <CHA>
A:Cross-references: UNIPROT:P01153
C:Superfamily: hypothalamic heptapeptide
C:Keywords: hypothalamus

Query Match      29.1%; Score 16; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 YHA 6
      |||
Db      3 YHS 5

RESULT 4
PT0285
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0285
R:Yanada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J
A:Reference number: PT0222; PMID:91108337; PMID:1899102
A:Accession: PT0285
A:Molecule type: DNA
A:Residues: 1-9 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match      29.1%; Score 16; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 PNYH 5
      |||
Db      4 PTYH 7

RESULT 5
S55696
C:Species: Trypanosoma brucei
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S55696
R:Hunt, M.; Kohler, P.
Biochim. Biophys. Acta 1249, 15-22, 1995
A:Title: Purification and characterization of phosphoenolpyruvate carboxykinase from Try
A:Reference number: S55696; PMID:95284106; PMID:7766679
A:Accession: S55696
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <HUN>
A:Cross-references: UNIPROT:Q7M3S5

Query Match      29.1%; Score 16; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 QPNYH 5
      |||
Db      1 QVNHQ 5

RESULT 7
S33844
C:Species: Bos primigenius taurus (cattle)
C:Date: 22-Nov-1993 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: S33844
R:Warburton, M.J.; Coles, B.; Dundas, S.R.; Gusterson, B.A.; O'Hare, M.J.
Eur. J. Biochem. 214, 803-809, 1993
A:Title: Hydrocortisone induces the synthesis of alpha(2)-macroglobulin by rat mammary m
A:Reference number: S33843; PMID:9330729; PMID:7686489
A:Accession: S33844
A:Molecule type: protein
A:Residues: 1-10 <WAR>
C:Superfamily: alpha-2-macroglobulin
C:Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond

Query Match      29.1%; Score 16; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.8e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 QPNYH 5
      |||
Db      1 QVNHQ 5

RESULT 8
PT0291
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0291
R:Yanada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J
A:Reference number: PT0222; PMID:91108337; PMID:1899102
A:Accession: PT0291
A:Molecule type: DNA
A:Residues: 1-10 <YAM>

```

A;Experimental source: B lymphocyte  
C;Keywords: heterotetramer; immunoglobulin

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYH 5  
DB 4 PYY 7

## RESULT 9

A59272  
peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase (EC 3.5.1.52) A, large chain -  
N;Alternate names: peptide N-glycosidase  
C;Species: Prunus dulcis var. saciva (sweet almond)  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C;Accession: A59272

R;Altman, F.; Paschinger, K.; Dalik, T.; Voraue, K.  
Eur. J. Biochem. 252, 118-123, 1998

A;Title: Characterisation of peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase A  
A;Reference number: A59272; MUID:98181894; PMID:9523720

A;Accession: A59272

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <ALT>

A;Cross-references: UNIPROT:P81899

C;Keywords: hydrolase

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 66.7%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHA 6  
DB 5 YHS 7

## RESULT 10

S23307  
neurokinin A - rainbow trout  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Aug-2004  
C;Accession: S23307

R;Jensen, J.; Conlon, J.M.

Eur. J. Biochem. 205, 659-664, 1992

A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod

A;Reference number: S23186; MUID:92298992; PMID:1376687

A;Accession: S23307

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <JEN>

A;Cross-references: UNIPROT:P28500

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 HAVN 8  
DB 1 HKIN 4

## RESULT 11

S23186  
neurokinin A - Atlantic cod  
C;Species: Gadus morhua (Atlantic cod)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 16-Aug-2004  
C;Accession: S23186

R;Jensen, J.; Conlon, J.M.

Eur. J. Biochem. 205, 659-664, 1992

A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod

A;Reference number: S23186; MUID:92298992; PMID:1376687

A;Accession: S23186

A;Molecule type: protein

A;Residues: 1-10 <JEN>

A;Cross-references: UNIPROT:P28500

A;Experimental source: brain

C;Function:

A;Description: may play a physiological role in the regulation of cardiovascular and gast

A;Note: neurokinin A is derived by post-translational processing of preprotachykinin A

C;Keywords: neuropeptide; amidated carboxyl end; tachykinin

F;10/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 29.1%; Score 16; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 HAVN 8  
DB 1 HKIN 4

## RESULT 12

A38081  
amine oxidase (copper-containing) (EC 1.4.3.6) - yeast (Pichia angusta) (fragment)  
C;Species: Pichia angusta  
C;Date: 31-Dec-1993 #sequence\_revision 03-Feb-1994 #text\_change 20-Apr-2000  
C;Accession: A38081

R;Mu, D.; Jones, S.M.; Smith, A.J.; Brown, D.E.; Dooley, D.M.; Klinman, J.P.

J. Biol. Chem. 267, 7979-7982, 1992

A;Title: Tyrosine codon corresponds to topa quinone at the active site of copper amine o>

A;Reference number: A38081; MUID:92235001; PMID:1569055

A;Accession: A38081

A;Molecule type: protein

A;Residues: 1-7 <MUA>

C;Keywords: copper; oxidoreductase; quinoxaline; topaquinone

F;4/Modified site: topaquinone (Tyr) #status experimental

Query Match 27.3%; Score 15; DB 2; Length 7;  
Best Local Similarity 60.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NYHAV 7  
DB 3 NYEV 7

## RESULT 13

A15398  
choline oxidase (EC 1.1.3.17) - Alcaligenes sp. (tentative sequence) (fragment)  
C;Species: Alcaligenes sp.  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C;Accession: A15398

R;Ohta-Fukuyama, M.; Miyake, Y.; Emi, S.; Yamano, T.

J. Biochem. 88, 197-203, 1980

A;Title: Identification and properties of the prosthetic group of choline oxidase from A]

A;Reference number: A15398; MUID:81006769; PMID:6997283

A;Accession: A15398

A;Molecule type: protein

A;Residues: 1-7 <OHT>

A;Cross-references: UNIPROT:P16101

C;Keywords: oxidoreductase

Query Match 27.3%; Score 15; DB 2; Length 7;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNY 4  
DB 3 PNH 5

## RESULT 14

A44960

neuropeptide Led-CC-I - Colorado potato beetle  
 C/Species: Leptinotarsa decemlineata (Colorado potato beetle)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C/Accession: A44960  
 R/Gaede, G.; Kellner, R.  
 Peptides 10, 1287-1289, 1989  
 A/Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and  
 A/Reference number: A44960; MUID:90160053; PMID:2576128  
 A/Accession: A44960  
 A/Molecule type: protein  
 A/Residues: 1-8 <GAE>  
 A/Cross-references: UNIPROT:P04548  
 C/Superfamily: adipokinetic hormone  
 C/Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 27.3%; Score 15; DB 2; Length 8;  
 Best Local Similarity 66.7%; Pred.No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNY 4  
 | |  
 | |  
 Db 6 PNW 8

## RESULT 15

B44960  
 neuropeptide Led-CC-II - Colorado potato beetle  
 C/Species: Leptinotarsa decemlineata (Colorado potato beetle)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C/Accession: B44960  
 R/Gaede, G.; Kellner, R.  
 Peptides 10, 1287-1289, 1989  
 A/Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and  
 A/Reference number: A44960; MUID:90160053; PMID:2576128  
 A/Accession: B44960  
 A/Molecule type: protein  
 A/Residues: 1-8 <GAE>  
 A/Cross-references: UNIPROT:P04549  
 C/Superfamily: adipokinetic hormone  
 C/Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
 F:8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 27.3%; Score 15; DB 2; Length 8;  
 Best Local Similarity 66.7%; Pred.No. 2.8e+05;  
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNY 4  
 | |  
 | |  
 Db 6 PNW 8

Search completed: October 27, 2004, 18:35:34  
 Job time : 29.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:53:09 ; Search time 124.5 seconds  
(without alignments)  
46.215 Million cell updates/sec

Title: US-09-867-159A-4  
Perfect score: 55  
Sequence: 1 QNYHAVNIV 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 2971

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_Q2:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	23	41.8	10	2	Q8J0C2	Q8J0C2 encephalito
2	20	36.4	10	2	Q9UCQ8	Q9UCQ8 homo sapien
3	19	34.5	9	2	P90359	P90359 barley mild
4	19	34.5	10	2	Q6R7V4	Q6R7V4 carlia zuma
5	19	34.5	10	2	AAS09890	AAS09890 carlia zu
6	18	32.7	9	1	BS43_SERPL	BS4375 seirratia pl
7	18	32.7	9	2	O08979	O08979 mus musculu
8	18	32.7	10	2	Q7RGV8	Q7RGV8 plasmodium
9	18	32.7	10	2	P82438	P82438 nicotiana t
10	17	30.9	7	1	FA34_PANRE	FA3438 panagrellus
11	17	30.9	8	1	LCR6_LEUMA	LCR618 leucophaea
12	17	30.9	10	2	Q9UCR0	Q9UCR0 homo sapien
13	17	30.9	10	2	Q7RSI4	Q7RSI4 plasmodium
14	17	30.9	10	2	H0QE18	H0QE18 human immun
15	16	29.1	7	1	HY7_PIG	P01153 sus scrofa
16	16	29.1	7	2	P72681	P72681 nocardia la
17	16	29.1	8	1	ALL1_CYDPO	P81152 cydia pomon
18	16	29.1	8	1	FA31_PANRE	P41872 panagrellus
19	16	29.1	8	2	Q9UDZ4	Q9UDZ4 homo sapien
20	16	29.1	8	2	Q9UNH9	Q9UNH9 homo sapien
21	16	29.1	8	2	P83158	P83158 anabaena sp
22	16	29.1	9	1	FA32_PANRE	P41873 panagrellus
23	16	29.1	9	2	Q7M3S5	Q7M3S5 trypanosoma
24	16	29.1	9	2	Q9TWX7	Q9TWX7 manduca sex
25	16	29.1	9	2	Q6XRV2	Q6XRV2 nectarinia
26	16	29.1	9	2	Q47410	Q47410 escherichia
27	16	29.1	9	2	Q9K4M6	Q9K4M6 staphylococ
28	16	29.1	9	2	AAP69808	AAP69808 nectarini
29	16	29.1	10	1	TKNB_ONCMY	P28500 oncorhynch
30	16	29.1	10	2	Q9UCQ4	Q9UCQ4 homo sapien
31	16	29.1	10	2	Q7M3I3	Q7M3I3 bos taurus

32	16	29.1	10	2	Q9TS43	Q9TS43 sus scrofa
33	16	29.1	10	2	Q6XFW8	Q6XFW8 aethopyga b
34	16	29.1	10	2	P81899	P81899 prunus dulc
35	16	29.1	10	2	Q86580	Q86580 simian para
36	16	29.1	10	2	AAP69782	AAP69782 aethopyga
37	15	27.3	7	1	CHOX_ALCSP	P16101 alcaligenes
38	15	27.3	8	1	AKH_MELML	P25423 melolontha
39	15	27.3	8	1	HTF1_PERAM	P04548 periplaneta
40	15	27.3	8	1	HTF2_PERAM	P04549 periplaneta
41	15	27.3	8	1	HTF_TENMO	Q15888 tenebrio mo
42	15	27.3	8	2	Q15888	Q15888 homo sapien
43	15	27.3	8	2	Q8IUB8	Q8IUB8 homo sapien
44	15	27.3	8	2	Q35792	Q35792 saccharomyc
45	15	27.3	8	2	Q6Y2F2	Q6Y2F2 citrus sine

ALIGNMENTS

RESULT 1  
Q8J0C2 PRELIMINARY; PRT; 10 AA.  
AC Q8J0C2;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE HSP70 (Fragment)  
OS Encephalitozoon hellem.  
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
OX NCBI\_TaxID=27973;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PV6/95, and PV7/95;  
RX MEDLINE=22838663; PubMed=12958242;  
RA Haro M., Del Aguila C., Fenoy S., Henriques-Gil N.;  
RT "Intraspecies genotype variability of the microsporidian parasite  
RT Encephalitozoon hellem."  
RL J. Clin. Microbiol. 41:4166-4171(2003).  
DR EMBL; AY171238; AAN73415.1; -  
DR EMBL; AY171239; AAN73417.1; -  
FT NON TER 10 10  
SQ SEQUENCE 10 AA; 1100 MW; 107C0D472DD44DD4 CRC64;

Query Match 41.8%; Score 23; DB 2; Length 10;  
Best Local Similarity 57.1%; Pred. No. 1e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNYHAVN 8  
||| :||  
Db 2 PNANALN 8

RESULT 2  
Q9UCQ8 PRELIMINARY; PRT; 10 AA.  
AC Q9UCQ8;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE AUTOTAXIN (Fragment)  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92129337; PubMed=1733949;  
RA Stracke M.L., Kruttsch H.C., Unsworth E.J., Cioce V.,  
RA Schifmann E., Liotta L.A.;  
RT "Identification, purification, and partial sequence analysis of  
RT autotaxin, a novel motility-stimulating protein."  
RL J. Biol. Chem. 267:2524-2529(1992).  
GO GO:0006928; P:cell motility; NAS.

```

FT NON_TER 1 1
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1171 MW; 736F44577AF1B2CB CRC64;

Query Match 36.4%; Score 20; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PNY 4
DB 7 PNY 9

RESULT 3
P90359 PRELIMINARY; PRT; 9 AA.
ID P90359;
AC P90359;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polypeptide (Fragment).
OS Barley mild mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC BYMOVIRUS.
OC NCBI_TaxID=12466;
RN [1]
RP SEQUENCE FROM N.A.
RA Peerboom E., Jacobs V., Cartwright E.J., Adams M.J., Steinbiss H.H.,
RA Antoniw J.F.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X96970; CAA65658.1; -.
KW Polyprotein.
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1177 MW; C40404473401F1B6 CRC64;

Query Match 34.5%; Score 19; DB 2; Length 9;
Best Local Similarity 42.9%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYHVN 8
DB 1 POTHRLN 7

RESULT 4
Q6R7V4 PRELIMINARY; PRT; 10 AA.
ID Q6R7V4;
AC Q6R7V4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN Name=GAPDH.
OS Carlia zuma.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Scincophora; Scincoidae;
OC Scincidae; Carlia.
OC NCBI_TaxID=260893;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolman G., Phillips B.;
RL "Single copy nuclear DNA markers characterized for comparative
phyllogeography in Australian wet tropics rainforest skinks.";
RT phyllogeography in Australian wet tropics rainforest skinks.";
RL Mol. Ecol. Notes 4:185-187(2004).
DR EMBL; AY508912; AAS09890.1; -.
FT NON_TER 1 1
SQ SEQUENCE 10 AA; 1171 MW; 9D0ABB2322C9C1EA CRC64;

Query Match 32.7%; Score 18; DB 1; Length 9;
Best Local Similarity 28.6%; Pred. No. 1.8e+06;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 PNYHVN 10
DB 1 POTHRLN 10
SQ SEQUENCE 10 AA; 1171 MW; 9D0ABB2322C9C1EA CRC64;

Query Match 34.5%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.5e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PNYHVN 10
DB 2 NDHFVKLV 9

RESULT 5
AAS09890 PRELIMINARY; PRT; 10 AA.
ID AAS09890;
AC AAS09890;
DT 12-MAY-2004 (TrEMBLrel. 27, Created)
DT 12-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 12-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN GAPDH.
OS Carlia zuma.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Scincophora; Scincoidae;
OC Scincidae; Carlia.
OC NCBI_TaxID=260893;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolman G., Phillips B.;
RL "Single copy nuclear DNA markers characterized for comparative
phyllogeography in Australian wet tropics rainforest skinks.";
RT phyllogeography in Australian wet tropics rainforest skinks.";
RL Mol. Ecol. Notes 4:185-187(2004).
DR EMBL; AY508912; AAS09890.1; -.
FT NON_TER 1 1
SQ SEQUENCE 10 AA; 1171 MW; 9D0ABB2322C9C1EA CRC64;

Query Match 34.5%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.5e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PNYHVN 10
DB 2 NDHFVKLV 9

RESULT 6
BS43_SERPL STANDARD; PRT; 9 AA.
ID BS43_SERPL;
AC P83375;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Bacteriocin serracin P 43 kDa subunit (Fragment).
OS Serratia plymuthica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OC NCBI_TaxID=82996;
RN [1]
RP SEQUENCE AND FUNCTION.
RC STRAIN=J7;
RC MEDLINE=22293561; PubMed=12406768;
RX Jabrane A., Sabri A., Compere P., Jacques P., Vandenberghe I.,
RA Van Beumen J., Thonart P.;
RT "Characterization of serracin P, a phage-tail-like bacteriocin, and
its activity against Erwinia amylovora, the fire blight pathogen.";
RL Appl. Environ. Microbiol. 68:5704-5710(2002).
CC -!- FUNCTION: Major component of a prophage tail sheath (Probable).
CC -!- FUNCTION: Antibacterial activity against gram-negative bacterium
E. amylovora.
CC Antibiotic; Bacteriocin; Direct protein sequencing.
KW NON_TER 9 9
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1095 MW; 1E66D412C871E1FB CRC64;

Query Match 32.7%; Score 18; DB 1; Length 9;
Best Local Similarity 28.6%; Pred. No. 1.8e+06;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 PNYHVN 10
DB 1 POTHRLN 10
SQ SEQUENCE 10 AA; 1171 MW; 9D0ABB2322C9C1EA CRC64;

Query Match 34.5%; Score 19; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 6.5e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```



```

Db      3 HHGVRL 9

RESULT 7
Q08979  PRELIMINARY; PRT; 9 AA.
ID      C08979
AC      C08979;
DT      01-JUL-1997 (TREMELrel. 04, Created)
DT      01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT      01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE      AML1 protein (Fragment)
GN      Name=AML1;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NMRI; TISSUE=Murine retrovirus induced tumor;
RX      MEDLINE=97332339; PubMed=9188573;
RA      Amtoft H.W., Sorensen A.B., Bareil C., Schmidt J., Luz A.,
RA      Pedersen F.S.;
RT      "Stability of AML1 (core) site enhancer mutations in T-lymphomas
RT      induced by attenuated S13-3 murine leukemia virus mutants.";
RL      J. Virol. 71:5080-5087(1997).
DR      EMBL; Y11802; CAA72496.1; -.
FT      NON TER 9 9
FT      NON TER 1 1
SQ      SEQUENCE 9 AA; 981 MW; 293E01E865A776D8 CRC64;

Query Match      32.7%; Score 18; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 QPNYH 5
      |||
      |||
      |||
DB      3 QPSGH 7

RESULT 8
Q7RGV8  PRELIMINARY; PRT; 10 AA.
ID      Q7RGV8;
AC      Q7RGV8;
DT      01-MAR-2004 (TREMELrel. 26, Created)
DT      01-MAR-2004 (TREMELrel. 26, Last sequence update)
DT      01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE      Hypothetical protein.
GN      Name=PY04238;
OS      Plasmodium yoelii yoelii.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=73239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=17XNL;
RX      PubMed=12368865;
RA      Carlton J.M., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA      Silva J.C., Ermolaeva M.D., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA      Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA      Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA      Cho J.K., Quackenbush J., Sedegah M., Shoab A., Cummings L.M.,
RA      Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris L.M.,
RA      Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA      van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA      Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA      Carucci D.J.;
RT      "Genome sequence and comparative analysis of the model rodent malaria
RT      parasite Plasmodium yoelii yoelii.";
RL      Nature 419:512-519(2002).
CC      -! CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
DR      EMBL; ABRL01001270; EAA16068.1; -.
CX      Hypothetical protein.

```

```

SQ      SEQUENCE 10 AA; 1332 MW; F8601A30545B5051 CRC64;

Query Match      32.7%; Score 18; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 1e+04;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      4 YHAVNI 9
      :|||
      :|||
      :|||
DB      2 HHYNI 7

RESULT 9
P82438  PRELIMINARY; PRT; 10 AA.
ID      P82438
AC      P82438;
DT      01-JUN-2000 (TREMELrel. 14, Created)
DT      01-JUN-2000 (TREMELrel. 14, Last sequence update)
DT      01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE      50 kDa cell wall protein (Fragment).
OS      Nicotiana tabacum (Common tobacco).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC      lamids; Solanales; Solanaceae; Nicotiana.
OX      NCBI_TaxID=4097;
RN      [1]
RP      SEQUENCE.
RC      STRAIN=cv. PETIT HAVANA;
RA      Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA      Wojtaszek P., Bolwell G.P.;
RT      "Proteomic study of secondary cell wall proteins from transformed
RT      tobacco culture.";
RL      Planta 0:0-0(2000).
CC      -! SUBCELLULAR LOCATION: CELL WALL.
CC      -! TISSUE SPECIFICITY: XYLEM.
DR      GO; GO:0005618; C:cell wall; IEA.
KW      Cell wall.
FT      NON TER 10 10
FT      NON TER 10 10
SQ      SEQUENCE 10 AA; 1126 MW; C68E32486AF77B46 CRC64;

Query Match      32.7%; Score 18; DB 2; Length 10;
Best Local Similarity 37.5%; Pred. No. 1e+04;
Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      2 PNYHAVNI 9
      |||
      |||
      |||
DB      2 PQYPXGNV 9

RESULT 10
FAR4 PANRE STANDARD; PRT; 7 AA.
ID      FAR4 PANRE
AC      P41875;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      FMRFamide-like neuropeptide PF4 (KPNFIRF-amide).
OS      Panagrellus redivivus.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC      Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX      NCBI_TaxID=6233;
RN      [1]
RP      SEQUENCE, AND SYNTHESIS.
RX      MEDLINE=95232026; PubMed=7716079;
RA      Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P., Thim L.,
RA      Kubiak T.M., Martin R.A., Geary T.G.;
RT      "Isolation and preliminary biological characterization of
RT      knpfrfamide, a novel FMRFamide-related peptide from the free-living
RT      nematode, Panagrellus redivivus.";
RL      Peptides 16:87-93(1995).
CC      -! FUNCTION: Myoactive; induces a rapid concentration-dependent
CC      muscle tension increase.
CC      -! SUBCELLULAR LOCATION: Secreted.
CC      -! SIMILARITY: Belongs to the FARP (FMRFamide related peptide)

```

CC family.  
 KW Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 7 7 Phenylalanine amide.  
 SQ SEQUENCE 7 AA; 921 MW; 69D40059C4576350 CRC64;

Query Match 30.9%; Score 17; DB 1; Length 7;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QPNV 4  
 : : :  
 DB 1 KPNF 4

## RESULT 11

LCCK6\_LEUMA STANDARD; PRT; 8 AA.  
 AC P19988;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Leucokinin VI (I-VI).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 OC Blaberidae; Leucophaea.  
 OC NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Head;  
 RX MEDLINE=87052651; PubMed=2877794;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure, and synthesis of leucokinin V and VI:  
 RT myotonic peptides of Leucophaea maderae.";  
 RL Comp. Biochem. Physiol. 89C:27-30(1987).  
 CC -1- FUNCTION: This cephalomyotropic peptide stimulates contractile  
 CC activity of cockroach prothodum (hindgut).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 DR PIR; JS0316;  
 KW Amidation; Direct protein sequencing; Neuropeptide;  
 KW Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 Pyrrolidone carboxylic acid.  
 FT MOD\_RES 8 8 Glycine amide.  
 SQ SEQUENCE 8 AA; 935 MW; 9D635B1E9D5A5A6 CRC64;

Query Match 30.9%; Score 17; DB 1; Length 8;  
 Best Local Similarity 33.3%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 QPNVHA 6  
 : : : : :  
 DB 1 QSSFSH 6

## RESULT 12

Q9UCRO PRELIMINARY; PRT; 10 AA.  
 AC Q9UCRO;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE AUTOTAXIN (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92129337; PubMed=1733949;  
 RA Stracke M.L., Krutzsch H.C., Unsworth E.J., Cioce V.,  
 RA Schiffmann E., Liotta L.A.;  
 FT "Identification, purification, and partial sequence analysis of  
 FT autotaxin, a novel motility-stimulating protein.";

RL J. Biol. Chem. 267:2524-2529(1992).  
 DR GO; GO:0005576; C:extracellular; IDA.  
 DR GO; GO:0030334; P:regulation of cell migration; IDA.  
 FT NON\_TER 1 1  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1193 MW; 3E01FB40544B19D7 CRC64;

Query Match 30.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 25.0%; Pred. No. 1.6e+04;  
 Matches 2; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 PNYHAVNI 9  
 : : : : :  
 DB 2 PPFENINL 9

## RESULT 13

Q7RSI4 PRELIMINARY; PRT; 10 AA.  
 ID Q7RSI4;  
 AC Q7RSI4;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 DE Hypothetical protein.  
 GN Name=PY00373;  
 OS Plasmodium yoelii yoelii.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OC NCBI\_TaxID=73239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=17XNL;  
 RX PubMed=12368865;  
 RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perrea M.,  
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,  
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,  
 RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,  
 RA Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,  
 RA Florens L., Yates F.R., III, Raine J.D., Sinden R.E., Harris M.A.,  
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Valcy A.B.,  
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,  
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 RA Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria  
 RT parasite Plasmodium yoelii yoelii.";  
 RL Nature 419:512-519(2002).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AABL01000104; EAA15335.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 10 AA; 1279 MW; 11EBCEB04B4B50 CRC64;

Query Match 30.9%; Score 17; DB 2; Length 10;  
 Best Local Similarity 14.3%; Pred. No. 1.6e+04;  
 Matches 1; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 4 YHAVNIV 10  
 : : : : :  
 DB 4 YVVISVI 10

## RESULT 14

Q8QE18 PRELIMINARY; PRT; 10 AA.  
 ID Q8QE18;  
 AC Q8QE18;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Truncated envelope glycoprotein (Fragment).  
 GN Name=env;  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11676;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22961413; PubMed=14601597;
RA Masharsky A.E., Klimov N.A., Kozlov A.P.;
RT "Molecular cloning and analysis of full-length genome of HIV type 1
RL strains prevalent in countries of the former Soviet Union.";
RL AIDS Res. Hum. Retroviruses 19:933-939(2003).
DR EMBL; AF413977; AAL78427.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
KW Envelope Protein.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1274 MW; A9FD6CCB544326D6 CRC64;

Query Match 30.9%; Score 17; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNY 4
Db | | |
4 QXNY 7

RESULT 15
HY7_PIG STANDARD; PRT; 7 AA.
AC F01153;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothalamic heptapeptide.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE AND SYNTHESIS.
RX MEDLINE=81213980; PubMed=6263778;
RA Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
RA Saifran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
RT "Isolation, structure and synthesis of a heptapeptide with in vitro
RT ACTH-releasing activity from porcine hypothalamus.";
RL Horm. Metab. Res. 13:228-232(1981).
DR PIR; A01417; NYPG7.
KW Direct protein sequencing.
SQ SEQUENCE 7 AA; 957 MW; 632B45B1FB5059A0 CRC64;

Query Match 29.1%; Score 16; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.8e+06;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHA 6
Db | | |
3 YHS 5

```

Search completed: October 27, 2004, 18:23:49  
 Job time : 126.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:51:34 ; Search time 112 Seconds

(without alignments)  
32.029 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 356729299 residues

Total number of hits satisfying chosen parameters: 444336

Minimum DB seq length: 0  
Maximum DB seq length: 10

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	10	5	AAO20570 Cysteine
2	55	100.0	10	5	ABR98535 Cysteine
3	51	92.7	9	6	ABU11122 House dus
4	37	67.3	9	6	ABU11108 House dus
5	30	54.5	10	2	AAI23224 Peptide d
6	28	50.9	8	4	AAE10569 Soybean p
7	28	50.9	9	8	ADE85720 Human Bph
8	27	49.1	8	8	ADK09727 Human pap
9	27	49.1	9	8	ADK09741 Human pap
10	27	49.1	9	8	ADK09742 Human pap
11	27	49.1	9	8	ADK09740 Human pap
12	27	49.1	10	8	ADK09760 Human pap
13	26	47.3	8	5	ABP46729 Human Bly
14	26	47.3	8	7	ADG97556 scFV VHCD
15	26	47.3	9	6	ABR25375 Human can
16	26	47.3	9	6	ABR25590 Human can
17	26	47.3	9	6	ABR24431 Human can
18	26	47.3	9	6	ABR25198 Human can
19	26	47.3	10	2	AAK96521 Hepatitis
20	26	47.3	10	6	ABR24938 Human can
21	26	47.3	10	6	ABR24674 Human can
22	26	47.3	10	6	ABR25279 Human can
23	26	47.3	10	6	ABR25474 Human can
24	26	47.3	10	6	ABR25523 Human can
25	26	47.3	10	6	ABR25760 Human can

26 26 47.3 10 6 ABR24521 Human can  
27 26 47.3 10 6 ABR25108 Human can  
28 26 47.3 10 6 ABR25710 Human can  
29 25 45.5 8 4 AAU68238 Human Bre  
30 25 45.5 8 6 ABB99503  
31 25 45.5 8 6 ABG76122 Scrambled  
32 25 45.5 9 2 AAR73741 Antigen f  
33 25 45.5 9 2 AAR77563 HIV-B35-3  
34 25 45.5 9 6 ABU11107 House dus  
35 25 45.5 9 6 ABR05547 Human can  
36 25 45.5 9 6 ABR05672 Human can  
37 25 45.5 9 6 ABR05873 Human can  
38 25 45.5 10 1 AAP82560 Guanine n  
39 25 45.5 10 6 ABR05590 Human can  
40 25 45.5 10 6 ABR05041 Human can  
41 25 45.5 10 6 ABR05844 Human can  
42 24 43.6 6 8 ADO37899 Binding p  
43 24 43.6 6 8 ADO37682 Binding p  
44 24 43.6 6 8 ADO28421 Capture s  
45 24 43.6 6 8 ADO28204 Capture s

#### ALIGNMENTS

RESULT 1  
AAO20570  
ID AAO20570 standard; peptide; 10 AA.  
XX AC AAO20570;  
XX DT 02-JAN-2003 (first entry)  
XX DE Cysteine protease epitope peptide region, SEQ ID No 4.  
XX KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopical eczema; epitope.  
XX OS Dermatophagoides pteronyssinus.  
XX WO200278736-A2.  
XX PD 10-OCT-2002.  
XX PF 28-MAR-2002; 2002WO-FR001098.  
XX PR 30-MAR-2001; 2001FR-00004370.  
PR 03-MAY-2001; 2001FR-00005929.  
XX PR 29-MAY-2001; 2001US-00867159.  
XX PA (ANTI-) ANTIALIS SARL.  
XX PT Loria E, Terrasse G, Trehin Y;  
XX WPI; 2002-750636/81.  
PT Antiallergic compositions containing an anti-histamine, a histamine  
PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
PT for the allergen.  
XX PS Claim 14; Page 11; 32pp; French.  
XX CC The invention relates to antiallergic compositions containing an anti-  
CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
CC isolated nucleic acid molecule that has at least one polynucleotide  
CC sequence coding for the allergen, together with a pharmaceutical carrier.  
CC The pharmaceutical composition of the invention is useful as a non-  
CC specific antiallergic treatment, and also useful in the treatment of  
CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
CC allergic and atopical eczema. This sequence represents a peptide of a  
CC cysteine protease epitope region relating to the antiallergic

```

CC compositions of the invention
XX
SQ Sequence 10 AA;

Query Match      100.0%; Score 55; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 1 QPNYHAVNIV 10
    |||||
    |||||

RESULT 2
ABB98535
ID ABB98535 standard; peptide; 10 AA.
XX
AC ABB98535;
XX
XX 13-DEC-2002 (first entry)
XX
DE Cysteine protease epitope #2.
XX
XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunotherapy; allergen; allergic hypersensitivity reaction;
KW allergic asthma; allergic rhinitis; allergic atopic eczema;
KW cysteine protease.
XX
XX OS Dermatophagoides pteronyssinus.
XX
PN FR2822709-A1.
XX
PD 04-OCT-2002.
XX
XX 03-MAY-2001; 2001FR-00005929.
XX
PR 30-MAR-2001; 2001FR-00004370.
XX
XX (ANTI-) ANTIALIS SARL.
XX
PI Loria E, Terrasse G, Trehin Y;
XX
DR WPI; 2002-735037/80.
XX
XX Antiallergic composition, useful for preventing and treating e.g. asthma,
PT rhinitis or eczema, containing at least two of allergen, antihistamine
PT and histamine synthesis inhibitor.
XX
PS Claim 8; Page 6; 33pp; French.
XX
XX The present invention relates to an antiallergic pharmaceutical
CC composition (I) comprising a pharmaceutical carrier containing an active
CC agent combination of at least two of: an allergen; an antihistamine; and
CC a histamine synthesis inhibitor. (I) is used for treating or preventing
CC allergic hypersensitivity reactions, especially allergic asthma, allergic
CC rhinitis or allergic atopic eczema, in babies, children or adults. The
CC present sequence is a peptide fragment (epitope) of cysteine protease
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the
CC invention
XX
SQ Sequence 10 AA;

Query Match      100.0%; Score 55; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10
Db 1 QPNYHAVNIV 10
    |||||
    |||||

RESULT 3
ABU11122
ID ABU11122 standard; peptide; 9 AA.
XX
AC ABU11122;
XX
DT 05-FEB-2003 (first entry)
XX
DE House dust mite Der p1 antigen peptide #45.
XX
KW House dust mite; Der p1 antigen; human CD8 cell epitope; allergy;
KW immune response; atopic patient; CD8+ T-cell epitope; antiallergic.
XX
XX Dermatophagoides pteronyssinus.
XX
PN WO200281512-A1.
XX
PD 17-OCT-2002.
XX
XX 03-APR-2002; 2002WO-GB001534.
XX
PR 06-APR-2001; 2001GB-00008752;
XX
XX (ISIS-) ISIS INNOVATION LTD.
XX
XX Ogg G, Seneviratne S;
XX
XX WPI; 2003-058499/05.
XX
XX New peptide fragments of the Der p1 antigen of the house dust mite
PT Dermatophagoides pteronyssinus contain a human CD8+ T cell epitope and
PT are useful to treat and prevent allergy to the major house dust mite
PT allergen.
XX
PS Disclosure; Page 32; 47pp; English.
XX
XX The present invention relates to house dust mite (Dermatophagoides
CC pteronyssinus) Der p1 antigen peptides containing human CD8 cell
CC epitopes. The peptides of the invention are useful in the treatment of
CC human or animal patients, particularly to raise an immune response to the
CC Der p1 antigen. They are useful in the treatment and prevention of
CC allergies to the major house dust mite antigen, and to monitor disease
CC activity in atopic patients. ABU11078-ABU11146 represent house dust mite
CC Der p1 antigen peptides containing CD8+ T-cell epitopes
XX
SQ Sequence 9 AA;

Query Match      92.7%; Score 51; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 9
Db 1 QPNYHAVNIV 9
    |||||
    |||||

RESULT 4
ABU11108
ID ABU11108 standard; peptide; 9 AA.
XX
AC ABU11108;
XX
DT 05-FEB-2003 (first entry)
XX
DE House dust mite Der p1 antigen peptide #31.
XX
KW House dust mite; Der p1 antigen; human CD8 cell epitope; allergy;
KW immune response; atopic patient; CD8+ T-cell epitope; antiallergic.
XX
XX Dermatophagoides pteronyssinus.
XX
PN WO200281512-A1.
XX
PD 17-OCT-2002.
XX

```

PF 03-APR-2002; 2002WO-GB001534.  
XX PR 06-APR-2001; 2001GB-00008752.  
XX PA (ISIS-) ISIS INNOVATION LTD.  
XX PI Ogg G, Seneviratne S;  
XX DR WPI; 2003-058499/05.  
XX PT New peptide fragments of the Der p1 antigen of the house dust mite  
XX PT Dermatophagoides pteronyssinus contain a human CD8+ T cell epitope and  
XX PT are useful to treat and prevent allergy to the major house dust mite  
XX PT allergen.  
XX PS Disclosure; Page 31; 47pp; English.  
XX CC The present invention relates to house dust mite (Dermatophagoides  
XX CC pteronyssinus) Der p1 antigen peptides containing human CD8 cell  
XX CC epitopes. The peptides of the invention are useful in the treatment of  
XX CC human or animal patients, particularly to raise an immune response to the  
XX CC Der p1 antigen. They are useful in the treatment and prevention of  
XX CC allergies to the major house dust mite antigen, and to monitor disease  
XX CC activity in atopic patients. ABU1.078-ABU1.1146 represent house dust mite  
XX CC Der p1 antigen peptides containing CD8+ T-cell epitopes  
XX CC  
XX CC Sequence 9 AA;  
XX  
XX Query Match 67.3%; Score 37; DB 6; Length 9;  
XX Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 4 YHAVNIV 10  
XX Db 1 YHAVNIV 7  
XX  
XX RESULT 5  
XX AAY23224  
XX ID AAY23224 standard; peptide; 10 AA.  
XX AC AAY23224;  
XX XX  
XX DT 25-AUG-1999 (first entry)  
XX XX  
XX DE Peptide derived from beta-1,2-xylosyltransferase.  
XX XX  
XX KW Beta-1,2-xylosyltransferase; beta 1,2-linked xylose; beta-linked mannose;  
XX KW N-linked oligosaccharide; storage glycoprotein; allergenicity; soybean.  
XX OS Glycine max.  
XX PN WO9929835-A1.  
XX PD 17-JUN-1999.  
XX PF 08-DEC-1998; 98WO-US026047.  
XX PR 08-DEC-1997; 97US-0067932P.  
XX PR 05-JAN-1998; 98US-0070418P.  
XX XX  
XX PA (UYAR-) UNIV ARKANSAS.  
XX PI Elbein AD, Bannon GA;  
XX XX  
XX DR WPI; 1999-385597/32.  
XX PT A Soybean-derived xylosyltransferase.  
XX PS Claim 3; Fig 4; 58pp; English.  
XX CC AAY23220-24 represent peptides released by beta-1,2-xylosyltransferase by  
XX CC Endo lys C digestion. The specification describes a plant-derived beta

CC 1,2-xylosyltransferase, where the enzyme adds a beta 1,2-linked xylose to  
CC the beta -linked mannose on the N-linked oligosaccharides of storage  
CC glycoproteins. Xylose units on N-linked oligosaccharides may play a  
CC critical role in allergenicity of plant derived glycoproteins and may  
CC also be important in regulating the structure of the oligosaccharide  
CC chains and the targeting of these proteins. Purification of a  
CC xylosyltransferase is useful in order to study its properties and  
CC specificities in the absence of interfering activities and possible  
CC inhibitors. The antibody can be used to determine the localization of the  
CC xylosyltransferase in suspension culture soybean cells and the  
CC distribution and level of the enzyme in plants as well as its levels at  
CC various stages of growth  
XX Sequence 10 AA;  
XX  
XX Query Match 54.5%; Score 30; DB 2; Length 10;  
XX Best Local Similarity 66.7%; Pred. No. 33;  
XX Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 4 YHAVNI 9  
XX Db 4 YHAINL 9  
XX  
XX RESULT 6  
XX AAE10569  
XX ID AAE10569 standard; peptide; 8 AA.  
XX AC AAE10569;  
XX XX  
XX DT 10-DEC-2001 (first entry)  
XX XX  
XX DE Soybean peptide 3 to analyse A. thaliana beta 1,2-xylosyltransferase.  
XX KW Beta 1,2-xylosyltransferase; transgenic plant; medicament; glycoprotein;  
XX KW soybean.  
XX OS Glycine max.  
XX PN WO200164901-A1.  
XX PD 07-SEP-2001.  
XX PF 02-MAR-2001; 2001WO-EP002352.  
XX PR 03-MAR-2000; 2000AT-00000355.  
XX PA (GLOE/) GLOESSL J.  
XX PI Gloessel J, Strasser R, Mucha J, Mach L, Altmann F, Wilson IB;  
XX PI Steinkellner H;  
XX XX  
XX DR WPI; 2001-582160/65.  
XX PT Novel DNA molecule encoding beta 1,2-xylosyltransferase, useful for  
XX PT producing transgenic plants and plant cells with increased efficiency in  
XX PT producing glycoproteins.  
XX PS Example 1; Fig 1; 62pp; English.  
XX CC The invention relates to Arabidopsis thaliana beta 1,2-xylosyltransferase  
XX CC plant protein and its cDNA molecule. Beta 1,2-xylosyltransferase protein  
XX CC nucleic acid (PNA) molecule is useful for producing transgenic plants and  
XX CC plant cells with increased efficiency in producing glycoproteins. The  
XX CC invention also relates to a method for producing recombinant human  
XX CC glycoproteins which is suitable for medical use. Beta 1,2-  
XX CC xylosyltransferase DNA is useful for immobilisation on DNA microarrays,  
XX CC e.g. for finding homologous sequences or for expression studies in plants  
XX CC or non-vertebrate animals. The invention is also useful for inactivation,  
XX CC suppression or over expression and production of beta 1,2-  
XX CC xylosyltransferase. The present sequence is soybean peptide 3 which is  
XX CC used for analysing Arabidopsis thaliana beta 1,2-xylosyltransferase

SQ Sequence 8 AA;  
 Query Match 50.9%; Score 28; DB 4; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVN 8  
 |||||  
 Db 4 YHAIN 8

RESULT 7  
 ADE85720  
 ID ADE85720 standard; peptide; 9 AA.  
 AC ADE85720;  
 XX 12-FEB-2004 (first entry)  
 DT DT  
 DE Human EphA2 antibody Eph099B-233.152 VH CDR3 SEQ ID NO:24.  
 XX cancer; hyperproliferative cell disease; EphA2 antibody;  
 KW EphA2 agonistic antibody; cytostatic; antiasthmatic; antiproliferative;  
 KW antiinflammatory; vasotropic; respiratory; gene therapy;  
 KW metastatic cancer; asthma; psoriasis; inflammatory bowel disease;  
 KW smooth muscle restenosis; endothelial restenosis; Crohn's disease;  
 KW chronic obstructive pulmonary disease; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003094859-A2.  
 XX 20-NOV-2003.  
 XX 12-MAY-2003; 2003WO-US015044.  
 XX 10-MAY-2002; 2002US-0379322P.  
 PR 14-OCT-2002; 2002US-0418213P.  
 PR 03-APR-2003; 2003US-0460507P.  
 XX  
 XX (WEDI-) MEDIMUNE INC.  
 PA  
 XX Kinch MS, Carles-Kinch K, Kiener P, Langermann S;  
 XX WPI; 2004-012002/01.  
 DR N-PSDB; ADE85728.  
 XX  
 PT Treating cancer or a non-cancer hyperproliferative cell disease (e.g.  
 PT asthma, psoriasis, inflammatory bowel disease or restenosis) in a patient  
 PT comprises administering to the patient a therapeutic amount of an EphA2  
 PT antibody.  
 XX  
 PS Claim 54; SEQ ID NO 24; 173pp; English.  
 XX  
 CC The present invention describes a method for treating cancer or a non-  
 CC cancer hyperproliferative cell disease or disorder in a patient, which  
 CC comprises administering to the patient a therapeutic amount of an EphA2  
 CC antibody (1) that is an EphA2 agonistic antibody, an EphA2 cancer cell  
 CC phenotype inhibiting antibody, an exposed EphA2 epitope antibody, or an  
 CC antibody that binds EphA2 with a K-off of less than 3 x 10<sup>-3</sup> s<sup>-1</sup>. Also  
 CC described: (1) a pharmaceutical composition comprising a therapeutic  
 CC amount of (1) and a pharmaceutical carrier; (2) a cell line that produces  
 CC (1); (3) a hybridoma deposited with the ATCC accession number PTA-4572,  
 CC PTA-4573 or PTA-4574; (4) an isolated nucleic acid comprising a  
 CC nucleotide sequence encoding a light chain variable domain or a heavy  
 CC chain variable domain of the EphA2 antibody; (5) a vector comprising the  
 CC nucleic acid described above; (6) a host cell comprising the vector; (7)  
 CC methods of identifying the EphA2 agonistic antibody or the EphA2 antibody  
 CC that inhibits a cancer cell phenotype, that kills cancer cells having a  
 CC cancer cell phenotype or that preferentially binds an EphA2 epitope  
 CC exposed on cancer cells; and (8) a method of diagnosing, prognosing or  
 CC monitoring the efficacy of therapy for cancer in a patient known to or  
 CC suspected to have cancer. (1) has cytostatic, antiasthmatic,

CC antiproliferative, antiinflammatory, vasotropic and respiratory activities,  
 CC and can be used in gene therapy. The composition and methods are useful  
 CC in managing, diagnosing, preventing or treating hyperproliferative cell  
 CC diseases (i.e. metastatic cancer) or non-cancer hyperproliferative cell  
 CC diseases or disorders, such as asthma, psoriasis, inflammatory bowel  
 CC disease, smooth muscle restenosis, endothelial restenosis, Crohn's  
 CC disease or chronic obstructive pulmonary disease. They may also be used  
 CC for monitoring the efficacy of therapy for cancer in a patient known to  
 CC or suspected to have cancer, and in screening for anti-cancer drugs. The  
 CC present sequence is used in the exemplification of the present invention.

XX  
 SQ Sequence 9 AA;  
 Query Match 50.9%; Score 28; DB 8; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYHAVN 8  
 |||||  
 Db 2 PRYHAMD 8

RESULT 8  
 ADK09727  
 ID ADK09727 standard; peptide; 8 AA.  
 AC ADK09727;  
 XX 06-MAY-2004 (first entry)  
 DT DT  
 DE Human papillomavirus peptide #1782.  
 XX pathogenic virus; alternative reading frame; antigenic determinant;  
 KW virucide; vaccine; therapeutic agent; infection; HPV.  
 XX Human papillomavirus.  
 OS  
 XX WO2004011650-A2.  
 PN 05-FEB-2004.  
 PD 24-JUL-2003; 2003WO-EP008112.  
 PF 24-JUL-2002; 2002AT-00001124.  
 PR 11-JUL-2003; 2003EP-00450171.  
 XX (INTE-) INTERCELL AG.  
 PA  
 XX Mattner F, Schmidt W, Habel A;  
 XX WPI; 2004-169243/16.  
 DR  
 XX New polypeptide encoded by an alternative reading frame of a pathogenic  
 XX virus comprising an antigenic determinant, useful for treating or  
 XX preventing an infection with the pathogenic virus.  
 PS Claim 18; Page 193; 220pp; English.  
 XX  
 CC This invention relates to a novel polypeptide encoded by an alternative  
 CC reading frame of a pathogenic virus, where the polypeptide starts with a  
 CC methionine amino acid residue, which comprises an antigenic determinant  
 CC and more than 7 amino acid residues. The invention may be useful for the  
 CC production of compounds with a virucide activity or the development of a  
 CC vaccine. The polypeptide or its fragments may be useful as a therapeutic  
 CC agent. It is also useful for the manufacture of a medicament for treating  
 CC or preventing an infection with the pathogenic virus. The present  
 CC sequence is that of a human papillomavirus (HPV) epitope peptide of the  
 CC invention.

XX  
 SQ Sequence 8 AA;  
 Query Match 49.1%; Score 27; DB 8; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;



Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYH 5  
Db 2 QPRYH 6

RESULT 9  
ADK09741  
ID ADK09741 standard; peptide; 9 AA.  
XX AC ADK09741;  
DT 06-MAY-2004 (first entry)  
XX DE  
XX DE Human papillomavirus peptide #1796.  
XX KW pathogenic virus; alternative reading frame; antigenic determinant;  
XX KW virucide; vaccine; therapeutic agent; infection; HPV.  
XX OS Human papillomavirus.  
XX PN WO2004011650-A2.  
XX PD 05-FEB-2004.  
XX PF 24-JUL-2003; 2003WO-EP008112.  
XX PR 24-JUL-2002; 2002AT-00001124.  
XX PR 11-JUL-2003; 2003EP-00450171.  
XX PA (INTE-) INTERCELL AG.  
XX PI Mattner F, Schmidt W, Habel A;  
XX PI WPI; 2004-169243/16.  
XX PT New polypeptide encoded by an alternative reading frame of a pathogenic  
XX PT virus comprising an antigenic determinant, useful for treating or  
XX PT preventing an infection with the pathogenic virus.  
XX PS Claim 18; Page 193; 220pp; English.  
XX CC This invention relates to a novel polypeptide encoded by an alternative  
XX CC reading frame of a pathogenic virus, where the polypeptide starts with a  
XX CC methionine amino acid residue, which comprises an antigenic determinant  
XX CC and more than 7 amino acid residues. The invention may be useful for the  
XX CC production of compounds with a virucide activity or the development of a  
XX CC vaccine. The polypeptide or its fragments may be useful as a therapeutic  
XX CC agent. It is also useful for the manufacture of a medicament for treating  
XX CC or preventing an infection with the pathogenic virus. The present  
XX CC sequence is that of a human papillomavirus (HPV) epitope peptide of the  
XX CC invention.  
XX SQ Sequence 9 AA;

Query Match 49.1%; Score 27; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYH 5  
Db 1 QPRYH 5

RESULT 10  
ADK09742  
ID ADK09742 standard; peptide; 9 AA.  
XX AC ADK09742;  
DT 06-MAY-2004 (first entry)  
XX DE  
XX DE Human papillomavirus peptide #1796.  
XX KW pathogenic virus; alternative reading frame; antigenic determinant;  
XX KW virucide; vaccine; therapeutic agent; infection; HPV.  
XX OS Human papillomavirus.  
XX PN WO2004011650-A2.  
XX PD 05-FEB-2004.  
XX PF 24-JUL-2003; 2003WO-EP008112.  
XX PR 24-JUL-2002; 2002AT-00001124.  
XX PR 11-JUL-2003; 2003EP-00450171.

DE Human papillomavirus peptide #1797.  
XX pathogenic virus; alternative reading frame; antigenic determinant;  
XX KW virucide; vaccine; therapeutic agent; infection; HPV.  
XX OS Human papillomavirus.  
XX PN WO2004011650-A2.  
XX PD 05-FEB-2004.  
XX PF 24-JUL-2003; 2003WO-EP008112.  
XX PR 24-JUL-2002; 2002AT-00001124.  
XX PR 11-JUL-2003; 2003EP-00450171.  
XX PA (INTE-) INTERCELL AG.  
XX PI Mattner F, Schmidt W, Habel A;  
XX PI WPI; 2004-169243/16.  
XX PT New polypeptide encoded by an alternative reading frame of a pathogenic  
XX PT virus comprising an antigenic determinant, useful for treating or  
XX PT preventing an infection with the pathogenic virus.  
XX PS Claim 18; Page 193; 220pp; English.  
XX CC This invention relates to a novel polypeptide encoded by an alternative  
XX CC reading frame of a pathogenic virus, where the polypeptide starts with a  
XX CC methionine amino acid residue, which comprises an antigenic determinant  
XX CC and more than 7 amino acid residues. The invention may be useful for the  
XX CC production of compounds with a virucide activity or the development of a  
XX CC vaccine. The polypeptide or its fragments may be useful as a therapeutic  
XX CC agent. It is also useful for the manufacture of a medicament for treating  
XX CC or preventing an infection with the pathogenic virus. The present  
XX CC sequence is that of a human papillomavirus (HPV) epitope peptide of the  
XX CC invention.  
XX SQ Sequence 9 AA;

Query Match 49.1%; Score 27; DB 8; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYH 5  
Db 1 QPRYH 5

RESULT 11  
ADK09740  
ID ADK09740 standard; peptide; 9 AA.  
XX AC ADK09740;  
XX DT 06-MAY-2004 (first entry)  
XX DE Human papillomavirus peptide #1795.  
XX KW pathogenic virus; alternative reading frame; antigenic determinant;  
XX KW virucide; vaccine; therapeutic agent; infection; HPV.  
XX OS Human papillomavirus.  
XX PN WO2004011650-A2.  
XX PD 05-FEB-2004.  
XX PF 24-JUL-2003; 2003WO-EP008112.  
XX PR 24-JUL-2002; 2002AT-00001124.  
XX PR 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.  
 XX  
 PI Mattner F, Schmidt W, Habel A;  
 XX  
 DR WPI; 2004-169243/16.  
 XX  
 PT New polypeptide encoded by an alternative reading frame of a pathogenic  
 PT virus comprising an antigenic determinant, useful for treating or  
 PT preventing an infection with the pathogenic virus.  
 XX  
 FS Claim 18; Page 193; 220pp; English.  
 XX  
 CC This invention relates to a novel polypeptide encoded by an alternative  
 CC reading frame of a pathogenic virus, where the polypeptide starts with a  
 CC methionine amino acid residue, which comprises an antigenic determinant  
 CC and more than 7 amino acid residues. The invention may be useful for the  
 CC production of compounds with a virucide activity or the development of a  
 CC vaccine. The polypeptide or its fragments may be useful as a therapeutic  
 CC agent. It is also useful for the manufacture of a medicament for treating  
 CC or preventing an infection with the pathogenic virus. The present  
 CC sequence is that of a human papillomavirus (HPV) epitope peptide of the  
 CC invention.  
 XX  
 SQ Sequence 9 AA;  
 XX  
 Query Match 49.1%; Score 27; DB 8; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QPNYH 5  
 DB 4 QPRYH 8  
 RESULT 12  
 ADK09760  
 ID ADK09760 standard; peptide; 10 AA.  
 XX  
 AC ADK09760;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human papillomavirus peptide #1815.  
 XX  
 KW pathogenic virus; alternative reading frame; antigenic determinant;  
 KW virucide; vaccine; therapeutic agent; infection; HPV.  
 XX  
 OS Human papillomavirus.  
 XX  
 PN W02004011650-A2.  
 XX  
 PD 05-FEB-2004.  
 XX  
 PF 24-JUL-2003; 2003WO-EP008112.  
 XX  
 PR 24-JUL-2002; 2002AT-00001124.  
 PR 11-JUL-2003; 2003EP-00450171.  
 XX  
 PA (INTE-) INTERCELL AG.  
 XX  
 PI Mattner F, Schmidt W, Habel A;  
 XX  
 DR WPI; 2004-169243/16.  
 XX  
 PT New polypeptide encoded by an alternative reading frame of a pathogenic  
 PT virus comprising an antigenic determinant, useful for treating or  
 PT preventing an infection with the pathogenic virus.  
 XX  
 FS Claim 18; Page 193; 220pp; English.  
 XX  
 CC This invention relates to a novel polypeptide encoded by an alternative  
 CC reading frame of a pathogenic virus, where the polypeptide starts with a

CC methionine amino acid residue, which comprises an antigenic determinant  
 CC and more than 7 amino acid residues. The invention may be useful for the  
 CC production of compounds with a virucide activity or the development of a  
 CC vaccine. The polypeptide or its fragments may be useful as a therapeutic  
 CC agent. It is also useful for the manufacture of a medicament for treating  
 CC or preventing an infection with the pathogenic virus. The present  
 CC sequence is that of a human papillomavirus (HPV) epitope peptide of the  
 CC invention.  
 XX  
 SQ Sequence 10 AA;  
 XX  
 Query Match 49.1%; Score 27; DB 8; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QPNYH 5  
 DB 2 QPRYH 6  
 RESULT 13  
 ABP46729  
 ID ABP46729 standard; peptide; 8 AA.  
 XX  
 AC ABP46729;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human BLYS binding scFv VH CDR3 SEQ ID 2740.  
 XX  
 KW BLYS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200202641-A1.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 15-JUN-2001; 2001WO-US019110.  
 XX  
 PR 16-JUN-2000; 2000US-0212210P.  
 PR 17-OCT-2000; 2000US-0240816P.  
 PR 16-MAR-2001; 2001US-0276248P.  
 PR 21-MAR-2001; 2001US-0277379P.  
 PR 25-MAY-2001; 2001US-0293499P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (CAMS-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
 XX  
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
 XX  
 DR WPI; 2002-114799/15.  
 XX  
 PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.  
 XX  
 PS Claim 2; Page 3042; 3148pp; English.  
 XX  
 CC This invention describes novel antibodies that immunospecifically bind to  
 CC B Lymphocyte Stimulator (BLYS) polypeptides. BLYS is a member of the  
 CC tumour necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antirheumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of BLYS. The antibodies bind to BLYS  
 CC and so may be used to detect and quantitate the presence of BLYS in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of BLYS. They may also be

CC administered to treat diseases associated with aberrant Blys expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 8 AA;  
Query Match 47.3%; Score 26; DB 5; Length 8;  
Best Local Similarity 37.5%; Pred. No. 1.7e+06;  
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PNYHAVNI 9  
|:|:| :|:  
Db 1 PSYHYMDV 8  
RESULT 14  
ADG97556  
ID ADG97556 standard; peptide; 8 AA.  
XX  
AC ADG97556;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE scFV VHCDR3 peptide that immunospecifically binds Blys SeqID 2740.  
XX  
KW antibody; B lymphocyte stimulator; Blys; tumour necrosis factor;  
KW B cell proliferation; differentiation; scFv; myasthenia gravis;  
KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
KW antinflammatory; antiasthmatic; antiallergic; cytostatic.  
XX  
OS Unidentified.  
XX  
PN WO2003055979-A2.  
XX  
PD 10-JUL-2003.  
XX  
PF 14-NOV-2002; 2002WO-US036496.  
XX  
PR 16-NOV-2001; 2001US-0311469P.  
PR 19-DEC-2001; 2001US-0340817P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;  
XX  
DR WPI; 2003-505530/47.  
XX  
PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
PT (Blys), useful for detecting and treating diseases or disorders e.g.  
PT rheumatoid arthritis, asthma and leukemia.  
XX  
PS Example 1; SEQ ID NO 2740; 394pp; English.  
XX  
CC This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scFvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukaemia, carcinoma and

CC lymphoma. Accordingly, they can be described as exhibiting various  
CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC peptide sequence is a single chain antibody variable heavy CDR3 peptide  
CC that immunospecifically binds Blys of the invention.  
XX  
SQ Sequence 8 AA;  
Query Match 47.3%; Score 26; DB 7; Length 8;  
Best Local Similarity 37.5%; Pred. No. 1.7e+06;  
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PNYHAVNI 9  
|:|:| :|:  
Db 1 PSYHYMDV 8  
RESULT 15  
ABR25375  
ID ABR25375 standard; peptide; 9 AA.  
XX  
AC ABR25375;  
XX  
DT 19-MAY-2003 (first entry)  
XX  
DE Human cancer-related protein 185P3C3 HLA peptide #1010.  
XX  
KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
KW human leukocyte antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO200283921-A2.  
XX  
PD 24-OCT-2002.  
XX  
PF 10-APR-2002; 2002WO-US011654.  
XX  
PR 10-APR-2001; 2001US-0282739P.  
PR 10-APR-2001; 2001US-0283112P.  
PR 25-APR-2001; 2001US-0286630P.  
XX  
PA (AGEN-) AGENSYS INC.  
XX  
PI Jakobovits A, Challita-Bid PM, Faris M, Ge W, Hubert RS;  
PI Morrison K, Morrison KK, Raitano AB;  
XX  
DR WPI; 2003-075555/07.  
XX  
PT New composition comprising a substance that modulates the structure of  
PT proteins and polynucleotides, useful for therapeutic, prognostic and  
PT diagnostic reagents for eliciting cellular or humoral immune response in  
PT cancer patients.  
XX  
PS Claim 13; Page 380; 1021pp; English.  
XX  
CC The present invention relates to novel human cancer-related genes and  
CC proteins (AB278120-AB278168 and ABR01789-ABR01861). The genes and  
CC proteins are useful for eliciting a humoral or cellular immune response.  
CC The genes are useful as probes and primers for the amplification and/or  
CC detection of genes, mRNAs or their fragments, as reagents for the  
CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
CC directing the expression of the protein, as tools for modulating or  
CC inhibiting the expression of genes and/or translation of transcripts, and  
CC as therapeutic agents. The proteins and peptides are useful as  
CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
CC from the invention  
XX  
SQ Sequence 9 AA;  
Query Match 47.3%; Score 26; DB 6; Length 9;  
Best Local Similarity 37.5%; Pred. No. 1.7e+06;

Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 PNYHAVNI 9  
|::|:  
Db 2 PDFHSENL 9

Search completed: October 27, 2004, 18:19:36  
Job time : 115 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:19:45 ; Search time 85.5 seconds

(without alignments)  
37.920 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 173299

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/PCR\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/PCRUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	10	10	US-09-867-159A-4
2	30	54.5	10	9	Sequence 4, Appli
3	30	54.5	10	14	Sequence 5, Appli
4	28	50.9	8	16	Sequence 2, Appli
5	28	50.9	9	15	Sequence 24, Appli
6	26	47.3	8	14	Sequence 2740, Ap
7	26	47.3	8	14	Sequence 2740, Ap
8	26	47.3	10	9	Sequence 204, App
9	25	45.5	8	10	Sequence 162, App
10	25	45.5	8	13	Sequence 4, Appli
11	25	45.5	8	16	Sequence 4, Appli
12	24	43.6	6	17	Sequence 742, App
13	24	43.6	6	17	Sequence 959, App

14	24	43.6	7	14	US-10-052-578-164	Sequence 164, App
15	24	43.6	7	14	US-10-052-578-210	Sequence 210, App
16	24	43.6	7	14	US-10-053-520-164	Sequence 164, App
17	24	43.6	7	14	US-10-053-520-210	Sequence 210, App
18	24	43.6	7	14	US-10-053-498B-164	Sequence 164, App
19	24	43.6	7	14	US-10-053-498B-210	Sequence 210, App
20	24	43.6	7	15	US-10-358-146A-22	Sequence 22, Appli
21	24	43.6	7	15	US-10-358-146A-68	Sequence 68, Appli
22	24	43.6	7	15	US-10-328-953-167	Sequence 167, App
23	24	43.6	7	15	US-10-328-953-213	Sequence 213, App
24	24	43.6	7	16	US-10-258-144-57	Sequence 57, Appli
25	24	43.6	7	16	US-10-258-144-103	Sequence 103, App
26	24	43.6	9	9	US-09-826-177-56	Sequence 56, Appli
27	24	43.6	9	15	US-10-428-335-142	Sequence 142, App
28	23	41.8	6	17	US-10-699-088-760	Sequence 760, App
29	23	41.8	6	17	US-10-699-088-1053	Sequence 1053, Ap
30	23	41.8	7	14	US-10-285-394-289	Sequence 289, App
31	23	41.8	7	15	US-10-601-837-176	Sequence 176, App
32	23	41.8	7	17	US-10-700-330-190	Sequence 190, App
33	23	41.8	9	9	US-09-826-177-40	Sequence 40, Appli
34	23	41.8	9	9	US-09-826-177-72	Sequence 72, Appli
35	23	41.8	9	9	US-09-826-177-74	Sequence 74, Appli
36	23	41.8	9	9	US-09-826-177-84	Sequence 84, Appli
37	23	41.8	9	9	US-10-254-446A-109	Sequence 109, App
38	23	41.8	9	16	US-10-415-014-106	Sequence 106, App
39	23	41.8	9	16	US-10-415-014-219	Sequence 219, App
40	23	41.8	9	16	US-10-415-014-414	Sequence 414, App
41	23	41.8	9	16	US-10-415-014-525	Sequence 525, App
42	23	41.8	9	16	US-10-415-014-632	Sequence 632, App
43	23	41.8	9	16	US-10-415-014-638	Sequence 638, App
44	23	41.8	10	14	US-10-082-014-90	Sequence 90, Appli
45	23	41.8				

ALIGNMENTS

RESULT 1  
US-09-867-159A-4  
; Sequence 4, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALSIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE OF INVENTION: and at least one anti-histamine compound  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(10)  
; OTHER INFORMATION: Compriss epitope from cystine protease.  
US-09-867-159A-4

Query Match 100.0%; Score 55; DB 10; Length 10;  
Best Local Similarity 100.0%; Pred. NO. 0.00-4;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPNYHAVNIV 10

Db 1 QPNYHAVNIV 10

RESULT 2  
 US-09-748-578-5  
 ; Sequence 5, Application US/09748578  
 ; Patent No. US20010016344A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Elbein, Alan D.  
 ; APPLICANT: Bannon, Gary A.  
 ; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof  
 ; FILE REFERENCE: D6063/D  
 ; CURRENT APPLICATION NUMBER: US/09/748,578  
 ; CURRENT FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: US 09/207,223  
 ; PRIOR FILING DATE: 1998-12-08  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SEQ ID NO 5  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: soybean  
 ; FEATURE:  
 ; OTHER INFORMATION: Amino acid sequence of a peptide released by Endo  
 ; OTHER INFORMATION: lys C digestion of purified xylosyltransferase.  
 US-09-748-578-5

Query Match 54.5%; Score 30; DB 9; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAWN 9  
 |||:  
 Db 4 YHAINL 9

RESULT 3  
 US-10-411-905-5  
 ; Sequence 5, Application US/10411905  
 ; Publication No. US20030166012A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Elbein, Alan D.  
 ; APPLICANT: Bannon, Gary A.  
 ; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof  
 ; FILE REFERENCE: D6063/D2  
 ; CURRENT APPLICATION NUMBER: US/10/411,905  
 ; CURRENT FILING DATE: 2003-04-11  
 ; PRIOR APPLICATION NUMBER: US 09/748,578  
 ; PRIOR FILING DATE: 2000-12-22  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SEQ ID NO 5  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: soybean  
 ; FEATURE:  
 ; OTHER INFORMATION: Amino acid sequence of a peptide released  
 ; OTHER INFORMATION: by Endo lys C digestion of purified  
 ; OTHER INFORMATION: xylosyltransferase.  
 US-10-411-905-5

Query Match 54.5%; Score 30; DB 14; Length 10;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAWN 9  
 |||:  
 Db 4 YHAINL 9

RESULT 4  
 US-10-220-467A-2  
 ; Sequence 2, Application US/10220467A  
 ; Publication No. US20040121325A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glossl Prof., Josef  
 ; TITLE OF INVENTION: Beta 1, 2-Xylosyltransferase-gene from Arabidopsis  
 ; FILE REFERENCE: SONN:01905

; CURRENT APPLICATION NUMBER: US/10/220,467A  
 ; CURRENT FILING DATE: 2003-12-22  
 ; PRIOR APPLICATION NUMBER: A 355/2000  
 ; PRIOR FILING DATE: 2000-03-03  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: soyabean  
 US-10-220-467A-2

Query Match 50.9%; Score 28; DB 16; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAWN 8  
 |||:  
 Db 4 YHAIN 8

RESULT 5  
 US-10-436-782-24  
 ; Sequence 24, Application US/10436782  
 ; Publication No. US20040028685A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kinch, Michael  
 ; APPLICANT: Carles-Kinch, Kelly  
 ; APPLICANT: Kiener, Peter  
 ; APPLICANT: Langermann, Solomon  
 ; TITLE OF INVENTION: EphA2 Monoclonal Antibodies and Methods of Use Thereof  
 ; FILE REFERENCE: 10271-097  
 ; CURRENT APPLICATION NUMBER: US/10/436,782  
 ; CURRENT FILING DATE: 2003-05-12  
 ; PRIOR APPLICATION NUMBER: 60/379,322  
 ; PRIOR FILING DATE: 2002-05-10  
 ; PRIOR APPLICATION NUMBER: 60/418,213  
 ; PRIOR FILING DATE: 2002-10-14  
 ; PRIOR APPLICATION NUMBER: 60/418,213  
 ; PRIOR FILING DATE: 2003-04-03  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: Patentin version 3.2  
 ; SEQ ID NO 24  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-436-782-24

Query Match 50.9%; Score 28; DB 15; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNYHAVN 8  
 |||:  
 Db 2 PRYHAMD 8

RESULT 6  
 US-09-880-748-2740  
 ; Sequence 2740, Application US/09880748  
 ; Publication No. US2003005937A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
 ; FILE REFERENCE: PFS23  
 ; CURRENT APPLICATION NUMBER: US/09/880,748  
 ; CURRENT FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: 60/212,210  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: 60/240,815  
 ; PRIOR FILING DATE: 2000-10-17  
 ; PRIOR APPLICATION NUMBER: 60/276,248  
 ; PRIOR FILING DATE: 2001-03-16

```
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2740
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2740

Query Match      47.3%; Score 26; DB 10; Length 8;
Best Local Similarity 37.5%; Pred. No. 1.2e+06;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 PNYHAVNI 9
DB      1 PSYHYMDV 8

RESULT 7
US-10-293-418-2740
; Sequence 2740, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2740
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2740

Query Match      47.3%; Score 26; DB 14; Length 8;
Best Local Similarity 37.5%; Pred. No. 1.2e+06;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 PNYHAVNI 9
DB      1 PSYHYMDV 8

RESULT 8
US-09-851-138-204
; Sequence 204, Application US/09851138
; Publication No. US20020183508A1
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
; AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
; AGENTS
```

```
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 6.0 / ASCII text output
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,138
; FILING DATE: 09-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,075
; FILING DATE: <unknown>
; APPLICATION NUMBER: EP 94870166.9
; FILING DATE: 21 Oct 1994
; APPLICATION NUMBER: EP 95870076.7
; FILING DATE: 28 Jun 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KAMMERER, PATRICIA A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: INNS:004
; INFORMATION FOR SEQ ID NO: 204:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 204:
US-09-851-138-204

Query Match      47.3%; Score 26; DB 9; Length 10;
Best Local Similarity 57.1%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 QPNYHAV 7
DB      1 RPKYHQV 7

RESULT 9
US-09-988-493-162
; Sequence 162, Application US/09988493
; Publication No. US20030064419A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
; APPLICANT: Page, Martin John
; APPLICANT: Parekh, Rajesh Shikhu
; APPLICANT: Waterfield, Michael Derek
; TITLE OF INVENTION: Proteins, Genes, and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Breast Cancer
; FILE REFERENCE: 2543-1-024
; CURRENT APPLICATION NUMBER: US/09/988,493
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: PCT/GB01/01219
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: GB 0006695.1
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: GB 0007265.2
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 8
; TYPE: PRT
; ORGANISM: homo sapien
US-09-988-493-162
```

Query Match 45.5%; Score 25; DB 10; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NYHAV 7  
 Db 2 NFHAV 6

## RESULT 10

US-10-007-363-4  
 ; Sequence 4, Application US/10007363  
 ; Publication No. US20020168354A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mochly-Rosen, Daria  
 ; TITLE OF INVENTION: pseudo-epsilon RACK Peptide Composition  
 ; TITLE OF INVENTION: and Method for Protection Against Tissue Damage Due to  
 ; TITLE OF INVENTION: Ischemia  
 ; FILE REFERENCE: 58600-8209.US00  
 ; CURRENT APPLICATION NUMBER: US/10/007,363  
 ; CURRENT FILING DATE: 2002-11-09  
 ; PRIOR APPLICATION NUMBER: US 60/247,830  
 ; PRIOR FILING DATE: 2000-11-10  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: scrambled pseudo-epsilon RACK octapeptide  
 US-10-007-363-4

Query Match 45.5%; Score 25; DB 13; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9  
 Db 1 PDYHDAGI 8

## RESULT 11

US-10-807-553-4  
 ; Sequence 4, Application US/10807553  
 ; Publication No. US20040186055A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mochly-Rosen, Daria  
 ; TITLE OF INVENTION: pseudo-epsilon RACK Peptide Composition  
 ; TITLE OF INVENTION: and Method for Protection Against Tissue Damage Due to  
 ; TITLE OF INVENTION: Ischemia  
 ; FILE REFERENCE: 58600-8209.US00  
 ; CURRENT APPLICATION NUMBER: US/10/807,553  
 ; CURRENT FILING DATE: 2004-03-22  
 ; PRIOR APPLICATION NUMBER: US/10/007,363  
 ; PRIOR FILING DATE: 2001-11-09  
 ; PRIOR APPLICATION NUMBER: US 60/247,830  
 ; PRIOR FILING DATE: 2000-11-10  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: scrambled pseudo-epsilon RACK octapeptide  
 US-10-807-553-4

Query Match 45.5%; Score 25; DB 16; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9  
 Db 1 PDYHDAGI 8

## RESULT 12

US-10-699-088-742  
 ; Sequence 742, Application US/10699088  
 ; Publication No. US20040209282A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dana Ault-Riche  
 ; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE  
 ; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES  
 ; FILE REFERENCE: 25885-1754  
 ; CURRENT APPLICATION NUMBER: US/10/699,088  
 ; CURRENT FILING DATE: 2003-10-30  
 ; PRIOR APPLICATION NUMBER: 60/422,923  
 ; PRIOR FILING DATE: 2002-10-30  
 ; PRIOR APPLICATION NUMBER: 60/423,018  
 ; PRIOR FILING DATE: 2002-10-30  
 ; NUMBER OF SEQ ID NOS: 1094  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 742  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic peptide  
 US-10-699-088-742

Query Match 43.6%; Score 24; DB 17; Length 6;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+06;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYH 5  
 Db 2 EPGYH 6

## RESULT 13

US-10-699-088-959  
 ; Sequence 959, Application US/10699088  
 ; Publication No. US20040209282A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dana Ault-Riche  
 ; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE  
 ; TITLE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES  
 ; FILE REFERENCE: 25885-1754  
 ; CURRENT APPLICATION NUMBER: US/10/699,088  
 ; CURRENT FILING DATE: 2003-10-30  
 ; PRIOR APPLICATION NUMBER: 60/422,923  
 ; PRIOR FILING DATE: 2002-10-30  
 ; PRIOR APPLICATION NUMBER: 60/423,018  
 ; PRIOR FILING DATE: 2002-10-30  
 ; NUMBER OF SEQ ID NOS: 1094  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 959  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic peptide  
 US-10-699-088-959

Query Match 43.6%; Score 24; DB 17; Length 6;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+06;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPNYH 5  
 Db 1 EPGYH 5



QY 2 PNYH 5  
| | : |  
Db 4 PNEH 7

Search completed: October 27, 2004, 18:39:36  
Job time : 86.5 secs

RESULT 14  
US-10-052-578-164  
; Sequence 164, Application US/10052578  
; Publication No. US20030134787A1  
; GENERAL INFORMATION:  
; APPLICANT: Sloan-Kettering Institute for Cancer Research  
; APPLICANT: Rothman, James E.  
; APPLICANT: Mayhew, Mark  
; APPLICANT: Hoe, Mee H.  
; APPLICANT: Houghton, Alan  
; APPLICANT: Hartl, Ulrich  
; APPLICANT: Querfelli, Ouathek  
; APPLICANT: Moroi, Yoichi  
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES  
; FILE REFERENCE: 11746/46003  
; CURRENT APPLICATION NUMBER: US/10/052,578  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 08/961,707  
; PRIOR FILING DATE: 1997-10-31  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 164  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide in m13 coliphage  
US-10-052-578-164

Query Match 43.6%; Score 24; DB 14; Length 7;  
Best Local Similarity 50.0%; Pred. No. 1.2e+06;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNYHAV 7  
| | | :  
Db 2 PRYHLI 7

RESULT 15  
US-10-052-578-210  
; Sequence 210, Application US/10052578  
; Publication No. US20030134787A1  
; GENERAL INFORMATION:  
; APPLICANT: Sloan-Kettering Institute for Cancer Research  
; APPLICANT: Rothman, James E.  
; APPLICANT: Mayhew, Mark  
; APPLICANT: Hoe, Mee H.  
; APPLICANT: Houghton, Alan  
; APPLICANT: Hartl, Ulrich  
; APPLICANT: Querfelli, Ouathek  
; APPLICANT: Moroi, Yoichi  
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES  
; FILE REFERENCE: 11746/46003  
; CURRENT APPLICATION NUMBER: US/10/052,578  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 08/961,707  
; PRIOR FILING DATE: 1997-10-31  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 210  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide in m13 coliphage  
US-10-052-578-210

Query Match 43.6%; Score 24; DB 14; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:16:05 ; Search time 27.5 Seconds  
(without alignments)  
24,116 Million cell updates/sec

Title: US-09-867-159A-4

Perfect score: 55

Sequence: 1 QPNYHAVNIV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 110780

Minimum DB seq length: 0

Maximum DB seq length: 10

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/protdata/1/aa/5A-COMB.pep:\*
  - 2: /cgn2\_6/protdata/1/aa/5B-COMB.pep:\*
  - 3: /cgn2\_6/protdata/1/aa/6A-COMB.pep:\*
  - 4: /cgn2\_6/protdata/1/aa/6B-COMB.pep:\*
  - 5: /cgn2\_6/protdata/1/aa/PCTUS-COMB.pep:\*
  - 6: /cgn2\_6/protdata/1/aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	54.5	10	3	US-09-207-223-5
2	30	54.5	10	4	US-09-748-578-5
3	26	47.3	10	3	US-08-836-075A-204
4	25	45.5	9	1	US-08-615-181-108
5	25	45.5	10	1	US-08-212-433A-33
6	25	45.5	10	3	US-08-716-256-33
7	25	45.5	10	5	PCT-US95-03239-33
8	25	45.5	10	6	5436320-3
9	23	41.8	6	4	US-08-877-605-181
10	23	41.8	6	4	US-08-877-605-220
11	22	40.0	4	4	US-09-057-363C-18
12	22	40.0	4	4	US-09-265-107-18
13	22	40.0	7	1	US-08-092-110A-3
14	22	40.0	7	1	US-08-273-474-3
15	22	40.0	7	3	US-08-405-647B-14
16	22	40.0	7	3	US-08-935-100-3
17	22	40.0	7	3	US-09-147-933-22
18	22	40.0	7	3	US-08-885-499-14
19	22	40.0	7	5	PCT-US94-07881-3
20	22	40.0	7	5	PCT-US96-03180-14
21	22	40.0	8	3	US-08-947-965-62
22	22	40.0	8	3	US-08-444-818-404
23	22	40.0	9	2	US-08-318-856A-34
24	22	40.0	10	3	US-08-836-075A-195
25	21	38.2	4	1	US-08-456-424-79
26	21	38.2	4	1	US-08-456-424-80
27	21	38.2	5	2	US-08-558-823-12

28	21	38.2	5	3	US-08-604-991-14	Sequence 14, Appl
29	21	38.2	5	3	US-09-363-639-14	Sequence 14, Appl
30	21	38.2	6	2	US-08-428-131-3	Sequence 3, Appl
31	21	38.2	6	2	US-08-558-823-15	Sequence 15, Appl
32	21	38.2	6	2	US-08-558-823-16	Sequence 16, Appl
33	21	38.2	6	2	US-08-310-912A-104	Sequence 104, Appl
34	21	38.2	6	3	US-08-893-534A-44	Sequence 44, Appl
35	21	38.2	6	3	US-08-841-089-104	Sequence 104, Appl
36	21	38.2	6	3	US-09-078-596-3	Sequence 3, Appl
37	21	38.2	6	3	US-08-996-679-44	Sequence 44, Appl
38	21	38.2	6	3	US-08-939-853A-28	Sequence 28, Appl
39	21	38.2	6	3	US-09-115-395-19	Sequence 19, Appl
40	21	38.2	6	3	US-09-301-085-104	Sequence 104, Appl
41	21	38.2	6	3	US-09-507-102-44	Sequence 44, Appl
42	21	38.2	6	3	US-09-250-059-34	Sequence 34, Appl
43	21	38.2	6	3	US-09-248-074-34	Sequence 34, Appl
44	21	38.2	6	3	US-09-187-859-60	Sequence 60, Appl
45	21	38.2	6	4	US-09-357-717-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-09-207-223-5  
; Sequence 5, Application US/09207223  
; Patent No. 6168937  
; GENERAL INFORMATION:  
; APPLICANT: Elbein, Alan D.  
; APPLICANT: Bannan, Gary A.  
; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof  
; FILE REFERENCE: D6063  
; CURRENT APPLICATION NUMBER: US/09/207,223  
; CURRENT FILING DATE: 1998-12-08  
; EARLIER APPLICATION NUMBER: US 60/067,932  
; EARLIER FILING DATE: 1997-12-08  
; NUMBER OF SEQ ID NOS: 7  
; SEQ ID NO 5  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: soybean  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence of a peptide released by Endo lys C  
; OTHER INFORMATION: digestion of purified xylosyltransferase.  
US-09-207-223-5

Query Match 54.5%; Score 30; DB 3; Length 10;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVNI 9  
| | | | |  
DB 4 YHAINL 9

RESULT 2  
US-09-748-578-5  
; Sequence 5, Application US/09748578  
; Patent No. 6593462  
; GENERAL INFORMATION:  
; APPLICANT: Elbein, Alan D.  
; APPLICANT: Bannan, Gary A.  
; TITLE OF INVENTION: Purified (1,2-Xylosyltransferase and Uses Thereof  
; FILE REFERENCE: D6063/D  
; CURRENT APPLICATION NUMBER: US/09/748,578  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/207,223  
; PRIOR FILING DATE: 1998-12-08  
; NUMBER OF SEQ ID NOS: 7  
; SEQ ID NO 5  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: soybean

```
;
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of a peptide released by Endo
; OTHER INFORMATION: lys C digestion of purified xylosyltransferase.
US-09-748-578-5

Query Match      54.5%; Score 30; DB 4; Length 10;
Best Local Similarity 56.7%; Pred. No. 14;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAWVI 9
   |||||
Db 4 YHAINL 9

RESULT 3
US-08-836-075A-204
; Sequence 204, Application US/08836075A
; Patent No. 6180768
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; TITLE OF INVENTION: STUTTER, LIEVEN
; TITLE OF INVENTION: NEW SEQUENCES OF HEPATITIS C VIRUS GENOTYPES
; TITLE OF INVENTION: AND THEIR USE AS PROPHYLACTIC, THERAPEUTIC AND DIAGNOSTIC
; TITLE OF INVENTION: AGENTS
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 6.0 / ASCII text output
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,075A
; FILING DATE: 21 Apr 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04155
; FILING DATE: 23 Oct 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94870166.9
; FILING DATE: 21 Oct 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95870076.7
; FILING DATE: 28 Jun 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KAMMERER, PATRICIA A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: INNS:004
; INFORMATION FOR SEQ ID NO: 204:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORGANISM: HUMAN
US-08-836-075A-204

Query Match      47.3%; Score 26; DB 3; Length 10;
Best Local Similarity 57.1%; Pred. No. 79;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QPNYHAV 7
   :|||
Db 1 RPXHQV 7

RESULT 4
US-08-615-181-108
; Sequence 108, Application US/08615181
```

```
; Patent No. 5756666
; GENERAL INFORMATION:
; APPLICANT: MASAFUMI, TAKIGUCHI
; APPLICANT: MIWA, KIYOSHI
; TITLE OF INVENTION: PEPTIDES CAPABLE OF INDUCING IMMUNE
; TITLE OF INVENTION: RESPONSE TO HIV AND ANTI-AIDS AGENT FOR PREVENTING AND
; TITLE OF INVENTION: CURING AIDS
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/615,181
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/01756
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 261302/1993
; FILING DATE: 19-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-796-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-08-615-181-108

Query Match      45.5%; Score 25; DB 1; Length 9;
Best Local Similarity 62.5%; Pred. No. 3.8e+05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PNYHAVNI 9
   |::||
Db 2 PSGHAVGI 9

RESULT 5
US-08-212-433A-33
; Sequence 33, Application US/08212433A
; Patent No. 5538897
; GENERAL INFORMATION:
; APPLICANT: Yates, III, John R.
; APPLICANT: Erg, James K.
; TITLE OF INVENTION: USE OF MASS SPECTROMETRY FRAGMENTATION
; TITLE OF INVENTION: PATTERNS OF PEPTIDES TO IDENTIFY AMINO ACID SEQUENCES IN
; TITLE OF INVENTION: DATABASES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart St. Tower
; CITY: San Francisco
```

STATE: CA  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/212,433A  
FILING DATE: 14-MAR-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Hughes, Richard L.  
REGISTRATION NUMBER: 31,264  
REFERENCE/DOCKET NUMBER: 16336-2  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-212-433A-33

Query Match 45.5%; Score 25; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4  
DB 6 QPNY 9

RESULT 6  
US-08-716-256-33  
Sequence 33, Application US/08/16256  
Patent No. 6017693  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: USE OF MASS SPECTROMETRY FRAGMENTATION  
PATTERNS TO IDENTIFY NUCLEOTIDE, AMINO ACID OR  
CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
NUMBER OF SEQUENCES: 46  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/716,256  
FILING DATE:  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA: PCT/US95/03239  
FILING DATE: 14-MAR-1995  
APPLICATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 16336-2PC  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-716-256-33

Query Match 45.5%; Score 25; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4  
DB 6 QPNY 9

RESULT 7  
PCT-US95-03239-33  
Sequence 33, Application PC/TUS9503239  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: USE OF MASS SPECTROMETRY FRAGMENTATION  
PATTERNS TO IDENTIFY NUCLEOTIDE, AMINO ACID OR  
CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
TITLE OF INVENTION: CARBOHYDRATE SEQUENCES IN DATABASES OR TO IDENTIFY  
NUMBER OF SEQUENCES: 46  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03239  
FILING DATE: 14-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,433  
FILING DATE: 14-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 16336-2PC  
TELEPHONE: 206-467-9600  
TELEFAX: 415-543-5043  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US95-03239-33

Query Match 45.5%; Score 25; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4  
DB 6 QPNY 9

RESULT 8  
5436320-3  
Patent No. 5436320  
APPLICANT: SPIEGEL, ALLEN M.  
TITLE OF INVENTION: ANTIBODY REAGENTS THAT IDENTIFY THE  
CARBOXY-TERMINAL PEPTIDE OF THE GTP-BINDING PROTEIN G  
NUMBER OF SEQUENCES: 10  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,377  
FILING DATE: 14-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 564,675

```
;
; FILING DATE: 08-AUG-1990
; APPLICATION NUMBER: 365,919
; FILING DATE: 15-JUN-1989
; APPLICATION NUMBER: 100,909
; FILING DATE: 25-SEP-1987
; SEQ ID NO:3
; LENGTH: 10
5436320-3

Query Match 45.5%; Score 25; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPNY 4
Db 6 QPNY 9

RESULT 9
US-08-877-605-181
; Sequence 181, Application US/08877605
; Patent No. 6582965
; GENERAL INFORMATION:
; APPLICANT: Robert Townsend
; APPLICANT: Raj Parekh
; APPLICANT: Sally Prime
; APPLICANT: Nick Webb
; TITLE OF INVENTION: A METHOD FOR DE NOVO PEPTIDE SEQUENCE DETERMINATION
; FILE REFERENCE: 9195-004
; CURRENT APPLICATION NUMBER: US/08/877,605
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 181
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide X Library
US-08-877-605-181

Query Match 41.8%; Score 23; DB 4; Length 6;
Best Local Similarity 60.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 YHAVN 8
Db 1 YHAID 5

RESULT 10
US-08-877-605-220
; Sequence 220, Application US/08877605
; Patent No. 6582965
; GENERAL INFORMATION:
; APPLICANT: Robert Townsend
; APPLICANT: Raj Parekh
; APPLICANT: Sally Prime
; APPLICANT: Nick Webb
; TITLE OF INVENTION: A METHOD FOR DE NOVO PEPTIDE SEQUENCE DETERMINATION
; FILE REFERENCE: 9195-004
; CURRENT APPLICATION NUMBER: US/08/877,605
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide X Library
US-08-877-605-220

;
; FILING DATE: 08-AUG-1990
; APPLICATION NUMBER: 365,919
; FILING DATE: 15-JUN-1989
; APPLICATION NUMBER: 100,909
; FILING DATE: 25-SEP-1987
; SEQ ID NO:3
; LENGTH: 10
5436320-3

Query Match 41.8%; Score 23; DB 4; Length 6;
Best Local Similarity 60.0%; Pred. No. 3.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 NYHAV 7
Db 1 DYHAI 5

RESULT 11
US-09-057-363C-18
; Sequence 18, Application US/09057363C
; Patent No. 6551994
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING THE
; INTERACTION BETWEEN ALPHA-CATENIN AND BETA-CATENIN
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,363C
; FILING DATE: 08-Apr-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 100086.406
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-057-363C-18

Query Match 40.0%; Score 22; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 HAVN 8
Db 1 HAVN 4

RESULT 12
US-09-265-107-18
; Sequence 18, Application US/09265107A
; Patent No. 6683048
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR STIMULATING
; GENE EXPRESSION AND CELLULAR DIFFERENTIATION
; FILE REFERENCE: 100086.406C1
; CURRENT APPLICATION NUMBER: US/09/265,107A
; CURRENT FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 75
```

```

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linear peptide modulating agent
US-09-265-107-18

Query Match          40.0%; Score 22; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 HAVN 8
Db      1 HAVN 4

RESULT 13
US-08-092-110A-3
; Sequence 3, Application US/08092110A
; Patent No. 5585477
; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David R.
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND
; TITLE OF INVENTION: METHODS OF DETECTION UTILIZING THE SAME
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Building, 127
; STREET: Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 13-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-273-474-3

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 13-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.075
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-092-110A-3

Query Match          40.0%; Score 22; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 NYHAVN 8
Db      2 NGHAIN 7

RESULT 14
US-08-273-474-3
; Sequence 3, Application US/08273474
; Patent No. 5691134
```

```

; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David R.
; TITLE OF INVENTION: POLIOVIRUS SPECIFIC PRIMERS AND METHODS
; TITLE OF INVENTION: OF DETECTION UTILIZING THE SAME
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Building, 127
; STREET: Peachtree Street, NE
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 13-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 688-0770
; TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-273-474-3

Query Match          40.0%; Score 22; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 3.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 NYHAVN 8
Db      2 NGHAIN 7

RESULT 15
US-08-405-647B-14
; Sequence 14, Application US/08405647B
; Patent No. 6124262
; GENERAL INFORMATION:
; APPLICANT: Sherman, Irwin W.
; APPLICANT: Crandall, Ian E.
; APPLICANT: Sholet, Stephen B.
; APPLICANT: Thevenin, Bernard Jean-Marie
; TITLE OF INVENTION: Compositions and Methods for Reducing
; TITLE OF INVENTION: Adhesiveness of Defective Red Blood Cells
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,647B
```

;  
; FILING DATE: 17-MAR-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 02307E-068700US  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-405-647B-14

Query Match 40.0%; Score 22; DB 3; Length 7;  
Best Local Similarity 75.0%; Pred. No. 3.8e+05;  
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNYH 5  
|  
|  
|  
Db 3 PKYH 6

Search completed: October 27, 2004, 18:36:37  
Job time : 28.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:12 ; Search time 2.40239 Seconds  
(without alignments)  
360.454 Million cell updates/sec

Title: US-09-867-159A-5

Perfect score: 57

Sequence: 1 WTVNSWDIT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	89.5	211	S21864	probable cysteine
2	51	89.5	245	JQ0337	allergen Der p 1 -
3	51	89.5	319	A61500	allergen Der f 1 p
4	48	84.2	184	S02729	actinidin (EC 3.4
5	48	84.2	380	1 TAGB	actinidin (EC 3.4
6	46	80.7	302	2 S02728	actinidin (EC 3.4
7	46	80.7	379	2 T37274	probable cathepsin
8	44	77.2	312	2 B23705	cysteine proteinase
9	44	77.2	312	2 A23705	cysteine proteinase
10	44	77.2	314	2 S41428	cysteine proteinase
11	44	77.2	325	2 S49451	cysteine proteinase
12	44	77.2	328	2 JQ1121	cysteine proteinase
13	44	77.2	355	2 T05390	probable cysteine
14	44	77.2	364	2 T46630	cysteine proteinase
15	44	77.2	364	2 T12039	cysteine proteinase
16	44	77.2	368	2 S47312	cysteine proteinase
17	44	77.2	374	2 T03941	cysteine proteinase
18	43	75.4	221	2 A53040	cysteine proteinase
19	43	75.4	326	2 S53027	cathepsin L (EC 3.
20	43	75.4	360	2 T08122	cysteine endopepti
21	43	75.4	371	2 JQ1111	cysteine proteinase
22	43	75.4	373	2 JQ1110	cysteine proteinase
23	42	73.7	328	2 T26446	hypothetical prote
24	42	73.7	378	2 T10949	cysteine proteinase
25	41	71.9	218	2 S67481	cathepsin L-like c
26	41	71.9	259	2 D89588	protein R09F10.1 (
27	41	71.9	324	2 S47432	cathepsin L (EC 3.
28	41	71.9	326	2 T09259	cathepsin L-like p
29	41	71.9	326	2 S43991	cathepsin L-like p

30 41 71.9 339 2 A53810  
31 41 71.9 343 2 D86198  
32 41 71.9 358 2 JC7787  
33 41 71.9 367 2 TC6529  
34 41 71.9 368 2 TC6529  
35 41 71.9 376 2 T03694  
36 41 71.9 376 2 T03694  
37 41 71.9 439 1 KHDOP  
38 41 71.9 463 2 S66504  
39 41 71.9 493 2 T01206  
40 41 71.9 658 2 TC8153  
41 40 70.2 73 2 A29172  
42 40 70.2 138 2 C84491  
43 40 70.2 139 2 C84491  
44 40 70.2 215 2 A59428  
45 40 70.2 294 2 S68784  
46 40 70.2 309 2 S41427

#### ALIGNMENTS

##### RESULT 1

S21864

probable cysteine proteinase (EC 3.4.22.-) - Euroglyphus maynei

N:Alternate names: allergen Eur m I

C:Species: Euroglyphus maynei

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C:Accession: S21864

R:Kent, N.A.; Hill, M.; Keen, J.N.; Holland, P.W.H.; Hart, B.

submitted to the EMBL Data Library, June 1991

A:Reference number: S21864

A:Accession: S21864

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-211 <KEN>

A:Cross-references: UNIPROT:P25780; EMBL:X60073

C:Genetics:

A:Introns: 100/3; 155/2

C:Superfamily: papain

C:Keywords: cysteine proteinase; hydrolase

Query Match 89.5%; Score 51; DB 2; Length 211;

Best Local Similarity 88.9%; Pred. No. 0.23;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WTVNSWDIT 9

Db 187 WTVNSWDIT 195

##### RESULT 2

JQ0337

allergen Der p 1 - house-dust mite (Dermatophagoides pteronyssinus) (fragment)

C:Species: Dermatophagoides pteronyssinus

C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C:Accession: JQ0337; A27582; A31657; C27634

R:Chua, K.Y.; Stewart, G.A.; Thomas, W.R.; Simpson, R.J.; Plozza, T.M.; J

J. Exp. Med. 167, 175-182, 1988

A:Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1.

A:Reference number: JQ0337; MUID:88089411; PMID:3335830

A:Accession: JQ0337

A:Molecule type: mRNA

A:Residues: 1-245 <CHU>

A:Cross-references: UNIPROT:P08176

R:Thomas, W.R.; Stewart, G.A.; Simpson, R.J.; Chua, K.Y.; Plozza, T.M.; Dilworth, R.J.; J

Int. Arch. Allergy Appl. Immunol. 85, 127-129, 1988

A:Title: Cloning and expression of DNA coding for the major house dust mite allergen Der

A:Reference number: A27582; MUID:88114080; PMID:3276629

A:Accession: A27582

A:Molecule type: mRNA

A:Residues: 6-101 <THO>

A:Cross-references: GB:M24794; NID:G387591; PIDN:AAA28296.1; PID:G387592

R:Simpson, R.J.; Nice, E.C.; Moritz, R.L.; Stewart, G.A.

Protein Seq. Data Anal. 2, 17-21, 1989  
 A>Title: Structural studies on the allergen Der p1 from the house dust mite Dermatophagoides pteronyssinus  
 A:Reference number: A31657; MUID:89098855; PMID:2911558  
 A:Accession: A31657

A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 24-56; 'XX', 59-64; 102-118; 134-149; 185-192, 'X', 194-200, 'X', 202 <SIM>  
 R:Lind, P.; Hansen, O.C.; Horn, N.  
 J. Immunol. 140, 4256-4262, 1988  
 A>Title: The binding of mouse hybridoma and human IgE antibodies to the major fecal allergen Der p1 by solid-phase inhibition assays with radiolabeled antigen.  
 A:Reference number: A92819; MUID:88229138; PMID:3372999

A:Accession: C27634  
 A:Molecule type: protein  
 A:Residues: 24-52 <LIN>  
 C:Superfamily: papain  
 C:Keywords: glycoprotein  
 F:24-245/Product: allergen Der p1 #status predicted <MAT>  
 F:75/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.5%; Score 51; DB 2; Length 245;  
 Best Local Similarity 88.9%; Pred. No. 0.27;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9  
 | | | | | | | |  
 Db 209 WIVRNSWDT 217

RESULT 3  
 A61500  
 allergen Der f I precursor - house-dust mite (Dermatophagoides farinae)  
 C:Species: Dermatophagoides farinae  
 C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 26-Aug-1999  
 C:Accession: A61500  
 R:Dilworth, R.J.; Chua, K.Y.; Thomas, W.R.  
 Clin. Exp. Allergy 21, 25-32, 1991  
 A>Title: Sequence analysis of cDNA coding for a major house dust mite allergen, Der f I.  
 A:Reference number: A61500; MUID:91215493; PMID:2021874  
 A:Accession: A61500

A>Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-319 <DIL>  
 C:Superfamily: papain

Query Match 89.5%; Score 51; DB 2; Length 319;  
 Best Local Similarity 88.9%; Pred. No. 0.35;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9  
 | | | | | | | |  
 Db 283 WIVRNSWDT 291

RESULT 4  
 S02729  
 actinidin (EC 3.4.22.14) precursor (clone pAC.7) - kiwi fruit (fragment)  
 C:Species: Actinidia chinensis (kiwi fruit)  
 C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 22-Jun-1999  
 C:Accession: S02729  
 R:Prækel, U.M.; McKee, R.A.; Smith, H.  
 Plant Mol. Biol. 10, 193-202, 1988

A>Title: Molecular analysis of actinidin, the cysteine proteinase of Actinidia chinensis  
 A:Reference number: S02728  
 A:Accession: S02729  
 A:Molecule type: mRNA  
 A:Residues: 1-184 <PRA>  
 A:Cross-references: EMBL:X13139; NID:g15958; PIDN:CAA31529.1; PID:g15959  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:1-159/Product: actinidin (fragment) #status experimental <MAT>

Query Match 84.2%; Score 48; DB 2; Length 184;

Best Local Similarity 77.8%; Pred. No. 0.63;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 WTVRNSWDT 9  
 | | | | | | | |  
 Db 117 WIVRNSWDT 125

# RESULT 5

TAGB  
 actinidin (EC 3.4.22.14) precursor - kiwi fruit  
 N:Alternate names: actinidin  
 C:Species: Actinidia chinensis var. deliciosa (kiwi fruit)  
 C:Date: 30-Nov-1980 #sequence\_revision 31-May-1996 #text\_change 18-Jun-1999  
 C:Accession: S12618; S12315; S12316; S06587; A00975  
 R:Snowden, K.C.; Gardner, R.C.  
 Nucleic Acids Res. 18, 6884, 1990  
 A>Title: Nucleotide sequence of an actinidin genomic clone.  
 A:Reference number: S12618; MUID:91067459; PMID:2251128  
 A:Accession: S12618  
 A:Molecule type: DNA  
 A:Residues: 1-380 <SNO>  
 A:Cross-references: EMBL:M38422; NID:gl66316; PIDN:AAA32629.1; PID:gl66317  
 R:Keeling, J.; Maxwell, P.; Gardner, R.C.  
 Plant Mol. Biol. 15, 787-788, 1990  
 A>Title: Nucleotide sequence of the promoter region from kiwifruit actinidin genes.  
 A:Reference number: S12315; MUID:91346716; PMID:2102886  
 A:Accession: S12315  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 81-100 <KEE>  
 A:Cross-references: EMBL:M35795  
 A:Accession: S12316  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-13 <KE2>  
 A:Cross-references: EMBL:M35795; NID:gl66322; PIDN:AAA32632.1; PID:g553033  
 R:Podivinsky, E.; Foster, R.L.S.; Gardner, R.C.  
 Nucleic Acids Res. 17, 8363, 1989  
 A>Title: Nucleotide sequence of actinidin, a kiwi fruit protease.  
 A:Reference number: S06587; MUID:90045955; PMID:2813065  
 A:Accession: S06587  
 A:Molecule type: mRNA  
 A:Residues: 1-122, 'F', 124-225, 'LD', 228-271, 'H', 273-349, 'H', 350, 'K', 352-372, 'D', 374-380 <F>  
 A:Cross-references: EMBL:X16466; NID:g15983; PIDN:CAA34486.1; PID:g15984  
 R:Carne, A.; Moore, C.H.  
 Biochem. J. 173, 73-83, 1978  
 A>Title: The amino acid sequence of the tryptic peptides from actinidin, a proteolytic enzyme from kiwifruit.  
 A:Reference number: A90300; MUID:78256777; PMID:687380  
 A:Accession: A00975  
 A:Molecule type: protein  
 A:Residues: 127-166, 'TS', 169, 'S', 171-191, 'D', 193-205, 'D', 207-211, 'Z', 213-222, 'D', 224, 'D',  
 A>Note: tryptic peptides were positioned on the basis of X-ray crystallographic data  
 C:Genetics:  
 A:Introns: 149/1; 228/3; 275/3; 364/3  
 C:Function:  
 A:Description: cysteine proteinase with broad specificity  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:1-24/Domains: signal sequence #status predicted <SIG>  
 F:25-126/Domains: amino-terminal propeptide #status predicted <PRO>  
 F:127-346/Product: actinidin #status experimental <MAT>  
 F:347-380/Domains: carboxyl-terminal propeptide #status predicted <CTP>  
 F:148-191, 192-224, 282-332/Disulfide bonds: #status experimental  
 F:151, 286, 345/Active site: Cys, His, Asn #status predicted

Query Match 84.2%; Score 48; DB 1; Length 380;  
 Best Local Similarity 77.8%; Pred. No. 1.3;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||:|||||  
 Db 304 WIVRNSWDT 312

RESULT 6  
 S02728  
 actinidain (EC 3.4.22.14) precursor (clone PAC.1) - kiwi fruit (fragment)  
 C:Species: Actinidia chinensis (kiwi fruit)  
 C>Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 04-Feb-2000  
 C:Accession: S02728  
 R:Prækel, U.M.; McKee, R.A.; Smith, H.  
 Plant Mol. Biol. 10, 193-202, 1988  
 A:Title: Molecular analysis of actinidin, the cysteine proteinase of Actinidia chinensis  
 A:Reference number: S02728  
 A:Accession: S02728  
 A:Molecule type: mRNA  
 A:Residues: 1-302 <PRA>  
 A:Cross-references: EMBL:X13013; NID:G15956; PIDN:CAA31435.1; PID:G15957  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:58-277/Product: actinidin #status experimental <MAT>  
 F:82,219,276/Active site: Cys, His, Asn #status predicted

Query Match 80.7%; Score 46; DB 2; Length 302;  
 Best Local Similarity 77.8%; Pred. No. 2.2;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||:|||||  
 Db 235 WIVRNSWDT 243

RESULT 7  
 T32774  
 probable cathepsin B (EC 3.4.22.1) cpr-6 - Caenorhabditis elegans  
 N:Alternate names: cathepsin B1  
 C:Species: Caenorhabditis elegans  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T32774; T37275; T34114  
 R:Larminie, C.G.; Johnstone, I.L.  
 DNA Cell Biol. 15, 75-82, 1996  
 A:Title: Isolation and characterization of four developmentally regulated cathepsin B-like genes from *Caenorhabditis elegans*  
 A:Reference number: Z21662; MUID:96152242; PMID:8561899  
 A:Accession: T32774  
 A:Molecule type: DNA  
 A:Residues: 1-379 <LAR>  
 A:Cross-references: UNIPROT:P43510; EMBL:L39939; PIDN:AAA98789.1  
 A:Experimental source: strain Bristol N2; clone c25b8  
 A:Accession: T32775  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-379 <LA2>  
 A:Cross-references: EMBL:L39894; NID:G671714; PIDN:AAA98787.1; PID:G671715  
 A:Experimental source: strain Bristol N2  
 R:Wilcox, L.  
 submitted to the EMBL Data Library, December 1995  
 A:Description: The sequence of C. elegans cosmid C25B8.  
 A:Reference number: Z21479  
 A:Accession: T34114  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-379 <WIL>  
 A:Cross-references: EMBL:U41556; PIDN:AACT0871.1; GSPDB:GN00028; CESP:C25B8.3  
 A:Experimental source: strain Bristol N2; clone C25B8  
 C:Genetics:  
 A:Gene: cpr-6; C25B8.3  
 A:Map position: X

A:Introns: 1/3; 27/1; 58/3; 90/3; 204/3; 295/3; 368/2  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:133,305,325/Active site: Cys, His, Asn #status predicted

Query Match 80.7%; Score 46; DB 2; Length 379;  
 Best Local Similarity 77.8%; Pred. No. 2.8;  
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||:|||||  
 Db 321 WTVRNSWDT 329

RESULT 8  
 B23705  
 cysteine proteinase (EC 3.4.22.-) - Entamoeba histolytica (strain SAW 1734) (fragment)  
 C:Species: Entamoeba histolytica  
 C>Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 04-Feb-2000  
 C:Accession: B23705  
 R:Tannich, E.; Scholze, H.; Nickel, R.; Horstmann, R.D.  
 J. Biol. Chem. 266, 4798-4803, 1991  
 A:Title: Homologous cysteine proteinases of pathogenic and nonpathogenic Entamoeba histolytica  
 A:Reference number: A23705; MUID:91161560; PMID:1705935  
 A:Accession: B23705  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-312 <TAN>  
 A:Cross-references: GB:M64721  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:115,256,276/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 312;  
 Best Local Similarity 77.8%; Pred. No. 5;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||:|||||  
 Db 272 WIVRNSWDT 280

RESULT 9  
 A23705  
 cysteine proteinase (EC 3.4.22.-) - Entamoeba histolytica (strain HM-1/IMSS) (fragment)  
 C:Species: Entamoeba histolytica  
 C>Date: 21-Feb-1992 #sequence\_revision 21-Feb-1992 #text\_change 09-Jul-2004  
 C:Accession: A23705  
 R:Tannich, E.; Scholze, H.; Nickel, R.; Horstmann, R.D.  
 J. Biol. Chem. 266, 4798-4803, 1991  
 A:Title: Homologous cysteine proteinases of pathogenic and nonpathogenic Entamoeba histolytica  
 A:Reference number: A23705; MUID:91161560; PMID:1705935  
 A:Accession: A23705  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-312 <TAN>  
 A:Cross-references: UNIPROT:Q01957; GB:M64721  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:115,256,276/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 312;  
 Best Local Similarity 77.8%; Pred. No. 5;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 |||:|||||  
 Db 272 WIVRNSWDT 280

RESULT 10  
 S41428  
 cysteine proteinase (EC 3.4.22.-) CP2 precursor - Trichomonas vaginalis

C:Species: Trichomonas vaginalis  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: S41428  
 R:Mailinsson, D.J.  
 Submitted to the EMBL Data Library, January 1994  
 A:Reference number: S41425  
 A:Accession: S41428  
 A:Molecule type: mRNA  
 A:Residues: 1-314 <MA>  
 A:Cross-references: UNIPROT:Q27107; EMBL:X77219; NID:9452293; PIDN:CAA54436.1; PID:94522  
 A:Experimental source: strain G3  
 C:Genetics:  
 A:Gene: CP2  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:1-98/Domain: signal sequence #status predicted <SIG>  
 F:99-314/Product: cysteine proteinase CP2 #status predicted <MAT>  
 F:123,260,280/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 314;  
 Best Local Similarity 77.8%; Pred. No. 5;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 276 WIVNSWGT 284

RESULT 11  
 S49451  
 Cysteine proteinase (EC 3.4.22.-) - chickpea  
 C:Species: Cicer arietinum (chickpea, garbanzo)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
 C:Accession: S49451; S31914; S46541  
 R:Cervantes, E.  
 Submitted to the EMBL Data Library, September 1994  
 A:Reference number: S49451  
 A:Accession: S49451  
 A:Molecule type: mRNA  
 A:Residues: 1-325 <CER>  
 A:Cross-references: UNIPROT:Q39451; EMBL:X82011; NID:9558562; PIDN:CAA57538.1; PID:95585  
 R:Cervantes, E.  
 Submitted to the EMBL Data Library, February 1993  
 A:Reference number: S31914  
 A:Accession: S31914  
 A:Molecule type: mRNA  
 A:Residues: 115-120, 'DVOP', 125, 'S', 126, 'F', 128-144, 'G', 146-209, 'NE', 211-223 <CE2>  
 A:Cross-references: EMBL:X70375; NID:922673; PIDN:CAA49836.1; PID:9536769  
 R:Cervantes, E.; Rodriguez, A.; Nicolas, G.  
 Plant Mol. Biol. 25, 207-215, 1994  
 A:Title: Ethylene regulates the expression of a cysteine proteinase gene during germination  
 A:Reference number: S46541; MUID:94289645; PMID:8018870  
 A:Accession: S46541  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 115-120, 'DVOP', 125, 'S', 126, 'F', 128-144, 'G', 146-210, 'EM', 212-223 <CE3>  
 A:Cross-references: EMBL:X70375; NID:922673  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:116,251,271/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 325;  
 Best Local Similarity 77.8%; Pred. No. 5.2;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 267 WIVNSWGT 275

RESULT 12  
 JQ1121  
 Cysteine proteinase (EC 3.4.22.-) COT44 [similarity] - rape

C:Species: Brassica napus (rape)  
 C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 09-Jul-2004  
 C:Accession: JQ1121  
 R:Dietrich, R.A.; Maslyar, D.J.; Heupel, R.C.; Harada, J.J.  
 Plant Cell 1, 73-80, 1989  
 A:Title: Spatial patterns of gene expression in Brassica napus seedlings: identification of proteinases.  
 A:Reference number: JQ1121; MUID:92386055; PMID:2535469  
 A:Accession: JQ1121  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <DIE>  
 A:Cross-references: UNIPROT:P25251  
 C:Comment: The mRNA encoding this protein is present in both cotyledons and axes, and this  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:121-163,155-196,254-305/Disulfide bonds: #status predicted  
 F:124,260,280/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 328;  
 Best Local Similarity 77.8%; Pred. No. 5.2;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 276 WIVNSWGT 284

RESULT 13  
 T05390  
 probable cysteine proteinase (EC 3.4.22.-) F16G20.220 - Arabidopsis thaliana  
 N:Alternate names: protein F16G20.220  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: T05390  
 R:Bevan, M.; Watson, M.D.; Gibbons, T.; Bartley, J.; Bancroft, I.; Mewes, H.W.; Mayer, K.  
 submitted to the Protein Sequence Database, August 1998  
 A:Reference number: Z15413  
 A:Accession: T05390  
 A:Molecule type: DNA  
 A:Residues: 1-355 <BEV>  
 A:Cross-references: UNIPROT:O81748; EMBL:AL031326; GSPDB:GN00062; ATSP:F16G20.220  
 A:Experimental source: cultivar Columbia; BAC clone F16G20  
 C:Genetics:  
 A:Gene: AtSP:F16G20.220  
 A:Map position: 4  
 A:Introns: 155/1; 237/3; 280/3  
 C:Superfamily: papain  
 C:Keywords: cysteine proteinase; hydrolase  
 F:157,293,313/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 355;  
 Best Local Similarity 77.8%; Pred. No. 5.7;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 309 WIVNSWGT 317

RESULT 14  
 T46630  
 cysteine proteinase (EC 3.4.22.-) 1 precursor [similarity] - kidney bean  
 N:Alternate names: cysteine endopeptidase 1  
 C:Species: Phaseolus vulgaris (kidney bean)  
 C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004  
 C:Accession: T46630  
 R:Sohlberg, L.E.; Sussex, I.M.  
 submitted to the EMBL Data Library, March 1996  
 A:Description: Sequence of a cDNA encoding a cysteine protease from germinating bean coty  
 A:Reference number: Z23106  
 A:Accession: T46630  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-364 <SOH>  
A:Cross-references: UNIPROT:Q41110; EMBL:U52970; NID:gi256829; PID:gi256830  
A:Experimental source: strain Taylor's Horticultural; cotyledon  
C:Genetics:  
A:Gene: CEP-1  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:149,285,305/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 364;  
Best Local Similarity 77.8%; Pred. No. 5.8;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSWD 9  
Db 301 WLVRSWGT 309

RESULT 15  
T12039  
cysteine proteinase (EC 3.4.22.-) 1 precursor - kidney bean  
C:Species: Phaseolus vulgaris (kidney bean)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T12039  
R:Senyuk, V.; Becker, C.; Muentz, K.  
A:Description: Isolation of cDNA clone encoding cysteine proteinase (CPI) from a cotyledon submitted to the EMBL Data Library, October 1997  
A:Reference number: Z17385  
A:Accession: T12039  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-364 <SEN>  
A:Cross-references: UNIPROT:O24321; EMBL:Z99952  
A:Experimental source: cultivar Moldavian; cotyledon; clone cp6a  
C:Superfamily: papain  
C:Keywords: cysteine proteinase; hydrolase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-124/Domain: propeptide #status predicted <PRO>  
F:125-364/Product: cysteine proteinase 1 #status predicted <MAT>  
F:149,285,305/Active site: Cys, His, Asn #status predicted

Query Match 77.2%; Score 44; DB 2; Length 364;  
Best Local Similarity 77.8%; Pred. No. 5.8;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSWD 9  
Db 301 WLVRSWGT 309

Search completed: October 27, 2004, 17:51:29  
Job time : 3.40239 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:34:27 ; Search time 13.3028 Seconds  
(without alignments)  
389.270 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	89.5	210	Q9GY0	Q9gy0 dermatophag
2	51	89.5	320	MMAL_DERPT	P08176 dermatophag
3	51	89.5	321	EUMI_BURMA	P25780 eurglyphus
4	51	89.5	321	MMAL_DERFA	P16311 dermatophag
5	51	89.5	321	BAC53948	Bac53948 dermatoph
6	50	87.7	235	Q9B116	Q9b116 paragonimus
7	50	87.7	325	Q46177	Q46177 paragonimus
8	48	84.2	184	Q96228	Q96228 actinidia d
9	48	84.2	217	Q9U0C5	Q9u0c5 clonorchis
10	48	84.2	320	Q9BKC0	Q9bkc0 clonorchis
11	48	84.2	327	Q86C54	Q86c54 clonorchis
12	48	84.2	380	ACTN_ACTCH	P00785 actinidia c
13	48	84.2	380	Q43367	Q43367 actinidia d
14	48	84.2	380	Q9AXD2	Q9axd2 actinidia c
15	46	80.7	302	Q96227	Q96227 actinidia d
16	46	80.7	357	Q9XF80	Q9xf80 sandersonia
17	46	80.7	360	Q7QP27	Q7qp27 giardia lam
18	46	80.7	382	Q7X750	Q7x750 glycine max
19	46	80.7	378	Q8MQC6	Q8mqc6 caenorhabdi
20	46	80.7	379	CPR6_CABEL	P43510 caenorhabdi
21	46	80.7	490	Q7XR52	Q7xr52 cryza sativ
22	44	77.2	45	Q9TXF1	Q9txf1 tritrichomo
23	44	77.2	188	Q24865	Q24865 entamoeba i
24	44	77.2	254	Q6URJ6	Q6urj6 trichomonas
25	44	77.2	254	AAR37419	Aar37419 trichomon
26	44	77.2	292	Q27109	Q27109 trichomonas
27	44	77.2	308	CPP3_ENTHI	Q06964 entamoeba h
28	44	77.2	310	ACP2_ENTHI	P36185 entamoeba h
29	44	77.2	314	Q27107	Q27107 trichomonas
30	44	77.2	315	CPPI_ENTHI	Q01957 entamoeba h
31	44	77.2	315	CPP2_ENTHI	Q01958 entamoeba h

32	44	77.2	318	2	Q95030	Q95030 entamoeba h
33	44	77.2	325	2	Q39451	Q39451 cicer ariet
34	44	77.2	328	1	CYS4_BRANA	P25251 brassica na
35	44	77.2	330	2	Q76852	Q76852 tetrahymena
36	44	77.2	348	2	Q84M27	Q84m27 helianthus
37	44	77.2	348	2	Q81748	Q81748 arabidopsis
38	44	77.2	355	2	Q81748	Q81748 arabidopsis
39	44	77.2	364	2	Q24321	Q24321 phaseolus v
40	44	77.2	364	2	Q41110	Q41110 phaseolus v
41	44	77.2	368	2	Q41696	Q41696 vicia sativ
42	44	77.2	374	2	Q24137	Q24137 nicotiana t
43	44	77.2	374	2	Q84YH7	Q84yh7 nicotiana t
44	43	75.4	111	2	Q46153	Q46153 penaeus van
45	43	75.4	111	2	Q7JMW6	Q7jmw6 penaeus van
	43	75.4	221	1	GPI_ZINOF	P82473 zingiber of

ALIGNMENTS

RESULT 1  
Q9GY0 PRELIMINARY; PRT; 210 AA.  
AC Q9GY0;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Allergen Der fi (Fragment).  
OS Dermatophagoides farinae (House-dust mite).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;  
OC Pyroglyphidae; Dermatophagoides.  
OX NCBI\_TaxID=6954;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hao M.Q., Xu J., Zhong N.S.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; AF285763; AAG00520.1; -.  
DR PIR; A27634; A27634.  
DR GO; GO:004197; F:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PR00705; PAPAIN.  
DR SMART; SM00645; Pept C1; 1.  
DR PROSITE; PS00640; THIOI\_PROTEASE ASN; 1.  
DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; UNKNOWN\_1.  
KW Hydrolase; Protease; Thiol protease.  
FT NON\_TER 1  
FT NON\_TER 210 210  
SQ SEQUENCE 210 AA; 23548 MW; BA08029D642EEB90 CRC64;  
Query Match 89.5%; Score 51; DB 2; Length 210;  
Best Local Similarity 88.9%; Pred. No. 0.84; 1; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 1;  
QY 1 WTVRNSWDT 9  
Db 186 WTVRNSWDT 194  
RESULT 2  
MMAL\_DERPT STANDARD; PRT; 320 AA.  
AC P08176; Q24616;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-FEB-1995 (Rel. 33, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Major mite fecal allergen Der p 1 precursor (EC 3.4.22.-) (Der p I).  
GN Name=DERP1;  
OS Dermatophagoides pteronyssinus (House-dust mite).

CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 CC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;  
 CC Pyroglyphidae; Dermatophagoides.  
 CC [1]  
 RN NCBI\_TaxID=6956;  
 CC  
 CC SEQUENCE FROM N.A., AND POLYMORPHISM.  
 RP MEDLINE=93357682; PubMed=8353459;  
 RA Chua K.Y., Kehal P.K., Thomas W.R.;  
 RT "Sequence polymorphisms of cDNA clones encoding the mite allergen Der  
 RT p 1.";  
 RL Int. Arch. Allergy Immunol. 101:364-368(1993).  
 RN [2]  
 CC SEQUENCE OF 76-320 FROM N.A.  
 RP MEDLINE=8809411; PubMed=333830;  
 RA Chua K.Y., Stewart G.A., Thomas W.R., Simpson R.J., Dilworth R.J.,  
 RA Plozza T.M., Turner K.J.;  
 RT "Sequence analysis of cDNA coding for a major house dust mite  
 RT allergen, Der p 1. Homology with cysteine proteases.";  
 RL J. Exp. Med. 167:175-182(1988).  
 RN [3]  
 CC SEQUENCE OF 81-176 FROM N.A.  
 RP MEDLINE=88114080; PubMed=3276629;  
 RA Thomas W.R., Stewart G.A., Simpson R.J., Chua K.Y., Plozza T.M.,  
 RA Dilworth R.J., Nisbet A., Turner K.J.;  
 RT "Cloning and expression of DNA coding for the major house dust mite  
 RT allergen Der p 1 in Escherichia coli.";  
 RL Int. Arch. Allergy Appl. Immunol. 85:127-129(1988).  
 RN [4]  
 CC REVISIONS TO 232-241.  
 RP MEDLINE=91215493; PubMed=2021874;  
 RA Dilworth R.J., Chua K.Y., Thomas W.R.;  
 RT "Sequence analysis of cDNA coding for a major house dust mite  
 RT allergen, Der p 1.";  
 RL Clin. Exp. Allergy 21:25-32(1991).  
 RN [5]  
 CC SEQUENCE OF 99-308 FROM N.A.  
 RP MEDLINE=91310112; PubMed=1483062;  
 RA Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;  
 RT "Molecular characterisation of group I allergen Eur m I from house  
 RT dust mite Euroglyphus maynei.";  
 RL Int. Arch. Allergy Immunol. 99:150-152(1992).  
 RN [6]  
 CC SEQUENCE OF 99-127.  
 RP MEDLINE=98229138; PubMed=3372999;  
 RA Lind P., Hansen O.C., Horn N.;  
 RT "The binding of mouse hybridoma and human IgE antibodies to the major  
 RT fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative  
 RT binding site location and species specificity studied by solid-phase  
 RT inhibition assays with radiolabeled antigen.";  
 RL J. Immunol. 140:4256-4262(1988).  
 RN [7]  
 CC SEQUENCE OF 99-139; 177-192; 208-224 AND 260-277, AND VARIANT ALA-222.  
 RP PubMed=2911558;  
 RA Simpson R.J., Nice E.C., Moritz R.L., Stewart G.A.;  
 RT "Structural studies on the allergen Der p 1 from the house dust mite  
 RT Dermatophagoides pteronyssinus: similarity with cysteine  
 RT proteinases.";  
 RL Protein Seq. Data Anal. 2:17-21(1989).  
 RN [8]  
 CC 3D-STRUCTURE MODELING.  
 RP MEDLINE=95062135; PubMed=7971950;  
 RA Topham C.M., Srinivasan N., Thorpe C.J., Overington J.P.,  
 RA Kalsheker N.A.;  
 RT "Comparative modelling of major house dust mite allergen Der p 1:  
 RT structure validation using an extended environmental amino acid  
 RT propensity table.";  
 RL Protein Eng. 7:869-894(1994).  
 RN [9]  
 CC -1- FUNCTION: Thiol protease that hydrolyzes proteins, with a  
 CC preference for phe or basic residues.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- ALLERGEN: Causes an allergic reaction in human. Common symptoms of  
 CC mite allergy are bronchial asthma, allergic rhinitis and  
 CC conjunctivitis. Reacts with IgE in 80% of patients with house dust

CC allergy.  
 CC -1- SIMILARITY: Belongs to peptidase family C1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U11695; AAB60215.1;  
 DR EMBL; M24794; AAA28296.1; ALT\_INIT.  
 DR EMBL; X65197; CAA46317.1;  
 DR PIR; JQ0337; JQ0337.  
 DR HSP; P53634; IK3B.  
 DR MEROPS; C01.073;  
 DR InterPro; IPR000668; Peptidase\_C1;  
 DR InterPro; IPR000169; Pept\_cys\_aseite.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPAIN\_1.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR SMART; SMO0645; Pept\_C1; 1.  
 DR PROSITE; PS00640; THIOI\_PROTEASE\_ASN; 1.  
 DR PROSITE; PS00139; THIOI\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOI\_PROTEASE\_HIS; 1.  
 DR Allergen; Direct protein sequencing; Glycoprotein; Hydrolase;  
 KW Polymorphism; Signal; Thiol protease; Zymogen.  
 FT SIGNAL 1 18 Potential  
 FT PROPEP 19 98 Activation peptide  
 FT CHAIN 99 320 Major mite fecal allergen Der p 1.  
 FT CARBOHYD 150 150 N-linked (GlcNAc...) (potential).  
 FT ACT\_SITE 132 132 By similarity.  
 FT ACT\_SITE 268 268 By similarity.  
 FT ACT\_SITE 288 288 By similarity.  
 FT DISULFID 102 215 By similarity.  
 FT DISULFID 129 169 By similarity.  
 FT DISULFID 163 201 By similarity.  
 FT VARIANT 148 148 Y -> H.  
 FT VARIANT 179 179 E -> K.  
 FT VARIANT 222 222 V -> A.  
 FT VARIANT 234 234 S -> T.  
 FT VARIANT 313 313 E -> Q.  
 SQ SEQUENCE 320 AA; 36104 MW; AOB1F4DD09791DFE CRC64;  
 Query Match 89.5%; Score 51; DB 1; Length 320;  
 Best Local Similarity 88.9%; Pred. No. 1.3;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 WTVNSWDT 9  
 Db 284 WIVNSWDT 292  
 RESULT 3  
 EUM1\_EURMA STANDARD; PRT; 321 AA.  
 AC P25780; Q9TZ23; Q9TZ24; Q9UBA0;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Mite group 1 allergen Eur m 1 precursor (EC 3.4.22.-) (Eur m I).  
 GN Name=EUM1;  
 OS Euroglyphus maynei (Mayne's house dust mite).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Acariformes; Sarcophormes; Astigmata; Psoroptidia; Analgoidea;  
 OC Pyroglyphidae; Euroglyphus.  
 OC NCBI\_TaxID=6958;  
 RN [1]  
 CC SEQUENCE FROM N.A. (EUR M 1.0101 AND EUR M 1.0102).  
 CC MEDLINE=99126275; PubMed=9925958;  
 CC Smith W., Mills K., Hazell L., Hart B.J., Thomas W.;  
 CC "Molecular analysis of the group 1 and 2 allergens from the house dust



```

RT mite, Euroglyphus maynei. ;
RL Int. Arch. Allergy Immunol. 118:15-22 (1999).
RN (2)
RP SEQUENCE OF 99-309 FROM N.A.
RX MEDLINE=93130112; PubMed=1483062;
RA Kent N.A., Hill M.R., Keen J.N., Holland P.W., Hart B.J.;
RT "Molecular characterisation of group I allergen Eur m I from house
RL Int. Arch. Allergy Immunol. 99:150-152 (1992).
CC -!- FUNCTION: Probable thiol protease.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of
CC mite allergy are bronchial asthma, allergic rhinitis and
CC conjunctivitis.
CC -!- SIMILARITY: Belongs to peptidase family C1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF047610; AAC82351.1; -
CC EMBL; AF047611; AAC82352.1; ALT INIT.
CC EMBL; AF047612; AAC82353.1; -
CC EMBL; X60073; CAA42677.1; -
CC PIR; S21864; S21864.
CC HSSP; P53634; 1K3B.
CC MEROPS; C01.073; -.
CC InterPro; IPR000668; Peptidase_C1.
CC InterPro; IPR000169; Pept_cys_acsite.
CC Pfam; PF00112; Peptidase_C1; 1.
CC PRINTS; PR00705; PAPA1N.
CC ProDom; PD000158; Peptidase_C1; 1.
CC SMART; SM00645; Pept_C1; 1.
CC PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
CC PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
CC PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
CC Allergen; Glycoprotein; Hydrolase; Signal; Thiol protease; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 98
FT CHAIN 99 321 Mite group I allergen Eur m 1.
FT ACT_SITE 133 133 By similarity.
FT ACT_SITE 269 269 By similarity.
FT ACT_SITE 289 289 By similarity.
FT ACT_SITE 130 170 By similarity.
FT DISULFID 34 34 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT VARIANT 36 126 T -> S (in Eur m 1.0102).
FT VARIANT 126 320 M -> N (in Eur m 1.0102).
FT VARIANT 320 320 M -> I (in Eur m 1.0102).
FT SEQUENCE 321 AA; 36290 MW; 6CFD44FEC725999E CRC64;
Query Match 89.5%; Score 51; DB 1; Length 321;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WTVRNSWDT 9
Db 285 WIVRNSWDT 293
RESULT 4
MMAL_DERFA STANDARD; PRT; 321 AA.
AC P16311;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Major mice fecal allergen Der f 1 precursor (EC 3.4.22.-) (Der f I).
GN Name=DERF1;

```

```

OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=91215493; PubMed=2021874;
RA Dilworth R.J., Chua K.Y., Thomas W.R.;
RT "The binding of mouse hyalidoma and human IgE antibodies to the major
RT fecal allergen, Der p 1, of Dermatophagoides pteronyssinus. Relative
RT binding site location and species specificity studied by solid-phase
RT inhibition assays with radiolabeled antigen.";
RL J. Immunol. 140:4256-4262 (1988)
CC -!- FUNCTION: Thiol protease that hydrolyzes proteins, with a
CC preference for Phe or basic residues.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human. Common symptoms of
CC mite allergy are bronchial asthma, allergic rhinitis and
CC conjunctivitis.
CC -!- SIMILARITY: Belongs to peptidase family C1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X65196; CAA46316.1; -
CC PIR; A27634; A27634.
CC HSSP; P53634; 1K3B.
CC MEROPS; C01.073; -.
CC InterPro; IPR000668; Peptidase_C1.
CC InterPro; IPR000169; Pept_cys_acsite.
CC Pfam; PF00112; Peptidase_C1; 1.
CC PRINTS; PR00705; PAPA1N.
CC ProDom; PD000158; Peptidase_C1; 1.
CC SMART; SM00645; Pept_C1; 1.
CC PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
CC PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
CC PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
CC Allergen; Direct protein sequencing; Glycoprotein; Hydrolase; Signal;
KW Thiol protease; Zymogen.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 98
FT CHAIN 99 321 Major mice fecal allergen Der f 1.
FT ACT_SITE 133 133 By similarity.
FT ACT_SITE 269 269 By similarity.
FT ACT_SITE 288 288 By similarity.
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT DISULFID 102 216 By similarity.
FT DISULFID 130 170 By similarity.
FT DISULFID 164 202 By similarity.
FT CONFLICT 201 201 R -> Q (in Ref. 2).
FT CONFLICT 282 282 D -> V (in Ref. 2).
FT SEQUENCE 321 AA; 36435 MW; 04523E54EBB476E CRC64;
Query Match 89.5%; Score 51; DB 1; Length 321;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 WTVNSWDT 9
Db      285 WIVNSWDT 293

RESULT 5
ID BAC53948 PRELIMINARY; PRT; 321 AA.
AC BAC53948;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DE 02-MAR-2004 (TRENBLrel. 27, Last annotation update)
DE Der f 1 allergen preproenzyme precursor.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotriiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Yasuhara T., Takai T., Takahashi K., Yuuki T., Yokota T., Okudaira H.,
RA Okumura Y.;
RT "Dermatophagoides farinae Der f 1 allergen preproenzyme mRNA.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Dilworth R.J., Chuan K.Y., Thomas W.R.;
RT "Sequence analysis of cDNA coding for a major house dust mite
RT allergen, Der f 1.";
RL Clin. Exp. Allergy 21:25-32(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93357682; PubMed=8353459;
RA Chua K.Y., Kehal P.K., Thomas W.R.;
RT "Sequence polymorphisms of cDNA clones encoding the mite allergen Der
RT p 1.";
RL Int. Arch. Allergy Immunol. 101:364-368(1993).
RN EMBL; AB034946; BAC53948.1; -.
KW SIGNAL.
FT CHAIN 1 18 POTENTIAL.
FT VARIANT 99 321
FT VARIANT 201 201 Q -> R (IN REF. 2).
FT VARIANT 282 282 V -> D (IN REF. 2).
SQ SEQUENCE 321 AA; 36391 MW; 83594754EBB4477 CRC64;

Query Match 89.5%; Score 51; DB 2; Length 321;
Best Local Similarity 88.9%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WTVNSWDT 9
Db      285 WIVNSWDT 293

RESULT 6
ID Q9BII6 PRELIMINARY; PRT; 235 AA.
AC Q9BII6;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Pre-procathepsin L.
OS Paragonimus westermani.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Plagiorchiida; Troglotremata; Paragonimidae; Paragonimus.
OX NCBI_TaxID=34504;
RN [1]
RP SEQUENCE FROM N.A.
RA Ling J., Zhang Y., Zhang Z.;
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family C1.
DR EMBL; AF362769; AAK35220.1; -.

QY      1 WTVNSWDT 9
Db      285 WIVNSWDT 293

RESULT 7
ID O46177 PRELIMINARY; PRT; 325 AA.
AC O46177;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Pre-procathepsin L precursor.
OS Paragonimus westermani.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Plagiorchiida; Troglotremata; Paragonimidae; Paragonimus.
OX NCBI_TaxID=34504;
RN [1]
RP SEQUENCE FROM N.A.
RA Yun D.H., Chung J.Y., Yang H.J., Cho S.Y., Kong Y.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family C1.
DR EMBL; U70537; AAB93494.1; -.
DR HSSP; P53634; 1K3B.
DR GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR SMART; SM00645; Pept C1; 1.
DR PROSITE; PS00640; THIOI_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; 1.
KW Hydrolase; Protease; Signal; Thiol protease.
FT SIGNAL 1 18 Potential.
FT CHAIN 112 325 cathepsin L.
SQ SEQUENCE 325 AA; 36722 MW; F3C18BF8A9ALEBAS CRC64;

Query Match 87.7%; Score 50; DB 2; Length 325;
Best Local Similarity 88.9%; Pred. No. 2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WTVNSWDT 9
Db      288 WTVNSWGT 296

RESULT 8
ID Q96228 PRELIMINARY; PRT; 184 AA.
AC Q96228;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)

```

DE Actinidin (Fragmest).  
OS Actinidia deliciosa (Kiwi).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Ericales; Actinidiaceae; Actinidia.  
OX NCBI\_taxid=3627;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Praekelt U.M., McKee R.A., Smith H.;  
RT "Molecular analysis of actinidin, the cysteine proteinase of actinidia  
chinesis".  
RL Plant Mol. Biol. 10:193-202(1988).  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; M21336; AA32631.1; -.  
DR GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase C1.  
DR InterPro; IPR000169; Pept\_cys\_acsite.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
KW Hydrolase; Protease; Thiol protease.  
FT NON TER 1  
FT CHAIN <1 159 actinidin.  
FT CHAIN 1  
SQ SEQUENCE 184 AA; 20109 MW; 52536AEC5D3F6AE CRC64;  
Query Match 84.2%; Score 48; DB 2; Length 184;  
Best Local Similarity 77.8%; Pred. No. 2.4;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WTVRNSWDT 9  
DB 117 WIVKNSWDT 125

RESULT 9  
ID Q9UOC5 PRELIMINARY; PRT; 217 AA.  
AC Q9UOC5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cysteine proteinase.  
GN Name=CSCP3;  
OS Clonorchis sinensis.  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
OC Opisthorchiida; Opisthorchiata; Opisthorchioidea; Opisthorchiidae;  
OC Clonorchis.  
OX NCBI\_taxid=79923;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Park H., Hong K.M., Ryu J.S., Shin C.H., Lee J.B., Soh C.T.,  
RA Paik M.K., Min D.Y.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; U85984; AF21471.1; -.  
DR HSP; Q9UBX1; 1M6D.  
DR GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR InterPro; IPR000169; Pept\_cys\_acsite.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PR00705; PAPAIN.  
DR ProDom; PD000158; Peptidase\_C1; 1.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
KW Hydrolase; Protease; Thiol protease.  
SQ SEQUENCE 217 AA; 23914 MW; CFF47D85654B04AE CRC64;  
Query Match 84.2%; Score 48; DB 2; Length 217;

Best Local Similarity 77.8%; Pred. No. 2.8;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WTVRNSWDT 9  
DB 180 WTVKNSWST 188

RESULT 10  
ID Q9BKCO PRELIMINARY; PRT; 320 AA.  
AC Q9BKCO;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cysteine proteinase 3.  
OS Clonorchis sinensis.  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
OC Opisthorchiida; Opisthorchiata; Opisthorchioidea; Opisthorchiidae;  
OC Clonorchis.  
OX NCBI\_taxid=79923;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yun D.-H., Chung J.-Y., Kong Y., Cho S.-Y.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family C1.  
DR EMBL; AF271091; AAK28439.1; -.  
DR HSP; P53634; IK3B.  
DR GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000668; Peptidase\_C1.  
DR InterPro; IPR000169; Pept\_cys\_acsite.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PR00705; PAPAIN.  
DR SMART; SM00645; Pept\_C1; 1.  
DR PROSITE; PS00640; THIOL\_PROTEASE\_ASN; 1.  
DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
KW Hydrolase; Protease; Thiol protease.  
FT CHAIN 114 320 cysteine proteinase3.  
FT CHAIN 114 320  
SQ SEQUENCE 320 AA; 36170 MW; BDED67FA14551D06 CRC64;  
Query Match 84.2%; Score 48; DB 2; Length 320;  
Best Local Similarity 77.8%; Pred. No. 4.3;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WTVRNSWDT 9  
DB 283 WTVKNSWST 291

RESULT 11  
ID Q86C54 PRELIMINARY; PRT; 327 AA.  
AC Q86C54;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Cysteine proteinase 3.  
OS Clonorchis sinensis.  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
OC Opisthorchiida; Opisthorchiata; Opisthorchioidea; Opisthorchiidae;  
OC Clonorchis.  
OX NCBI\_taxid=79923;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Nagano I., Pei F., Wu Z., Wu J., Cui H., Boonmars T., Takahashi Y.;  
RT "Molecular expression of a cysteine proteinase of Clonorchis sinensis  
and its application to an enzyme-linked immunosorbent assay for  
immunodiagnosis of clonorchiasis".  
RT Clin. Diagn. Lab. Immunol. 11:411-416 (2004).  
RL Clin. Diagn. Lab. Immunol. 11:411-416 (2004).  
CC -!- SIMILARITY: Belongs to peptidase family C1.

```

DR EMBL; AY273803; AAP33050.1; -.
DR HSP; Q9UBX1; IM6D.
DR GO; GO:0004137; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000668; Peptidase_C1.
DR InterPro; IPR000169; Pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPA1N.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SMO0645; Pept_C1; 1.
DR PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Protease; Thiol protease.
SQ SEQUENCE 327 AA; 37038 MW; 5F7D417290A2134D CRC64;

Query Match      84.2%; Score 48; DB 2; Length 327;
Best Local Similarity 77.8%; Pred. No. 4.4;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVKNSWDT 9
Db 290 WTVKNSWST 298

RESULT 12
ID ACTN_ACTCH STANDARD; PRT; 380 AA.
AC P00785;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Actinidin precursor (EC 3.4.22.14) (Actinidin) (Allergen Act c 1).
OS Actinidia chinensis (Kiwi) (Vangataol).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Ericales; Actinidiaceae; Actinidia.
OX NCBI_TaxID=3625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Deliciosa;
RX MEDLINE=90045955; PubMed=2813065;
RA Podivinsky E., Forster R.L.S., Gardner R.C.;
RT "Nucleotide sequence of actinidin, a kiwi fruit protease.";
RL Nucleic Acids Res. 17:8363-8363(1989).
RN [2]
RP SEQUENCE OF 70-371 FROM N.A.
RA Praekelt U.M., McKee R.A., Smith H.;
RT "Molecular analysis of actinidin, the cysteine proteinase of Actinidia chinensis.";
RL Plant Mol. Biol. 10:193-202(1988).
RN [3]
RP SEQUENCE OF 127-346.
RX MEDLINE=78256777; PubMed=687380;
RA Carne A., Moore C.H.;
RT "The amino acid sequence of the tryptic peptides from actinidin, a proteolytic enzyme from the fruit of Actinidia chinensis.";
RL Biochem. J. 173:73-83(1978).
RN [4]
RP SEQUENCE OF 1-13 FROM N.A.
RX MEDLINE=91346716; PubMed=2102886;
RA Keeling J., Maxwell P., Gardner R.C.;
RT "Nucleotide sequence of the promoter region from kiwifruit actinidin genes.";
RL Plant Mol. Biol. 15:787-788(1990).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS), AND REVISIONS.
RX MEDLINE=81072298; PubMed=7003158;
RA Baker E.N.;
RT "Structure of actinidin, after refinement at 1.7-A resolution.";
RL J. Mol. Biol. 141:441-484(1980).
CC -!- CATALYTIC ACTIVITY: Specificity close to that of papain.
CC -!- ALLERGEN: Causes an allergic reaction in human.

```

```

CC -!- SIMILARITY: Belongs to peptidase family C1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X16466; CAA34486.1; -.
CC EMBL; X13013; CAA31435.1; -.
CC EMBL; X13139; CAA31529.1; -.
CC EMBL; X57551; CAA40778.1; -.
CC PDB; 1AEC; X-ray; @=127-344.
CC PDB; 1F3Q; Model; A=127-344.
CC PDB; 1FOG; Model; A=127-344.
CC PDB; 2ACT; X-ray; @=127-346.
CC MEROPS; C01.007; -.
CC InterPro; IPR000658; Peptidase_C1.
CC InterPro; IPR000169; Pept_cys_acsite.
CC Pfam; PF00112; Peptidase_C1; 1.
CC PRINTS; PR00705; PAPA1N.
CC ProDom; PD000158; Peptidase_C1; 1.
CC SMART; SMO0645; Pept_C1; 1.
CC PROSITE; PS00640; THIOL_PROTEASE ASN; 1.
CC PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
CC PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW 3D-structure; Allergen; Direct protein sequencing; Hydrolase; Signal;
KW Thiol protease; Zymogen.
FT SIGNAL 1 24 Potential.
FT PROPEP 25 126 Activation peptide.
FT CHAIN 127 380 Actinidin.
FT ACT_SITE 151 151
FT ACT_SITE 288 288
FT ACT_SITE 308 308
FT DISULFID 148 191
FT DISULFID 182 224
FT DISULFID 282 332
FT DISULFID 96 96
FT CONFLICT 108 108
FT CONFLICT 123 124
FT CONFLICT 167 170
FT CONFLICT 181 181
FT CONFLICT 184 184
FT CONFLICT 192 192
FT CONFLICT 206 206
FT CONFLICT 212 212
FT CONFLICT 223 227
FT CONFLICT 230 231
FT CONFLICT 240 240
FT CONFLICT 272 274
FT CONFLICT 286 286
FT CONFLICT 290 291
FT CONFLICT 301 301
FT CONFLICT 307 307
FT CONFLICT 349 349
FT CONFLICT 360 360
FT CONFLICT 371 380
FT STRAND 131 132
FT HELIX 133 136
FT TURN 137 137
FT TURN 144 144
FT TURN 146 147
FT HELIX 151 168
FT STRAND 174 174
FT HELIX 176 182
FT STRAND 185 185
FT TURN 186 187
FT STRAND 188 188
FT HELIX 190 192
FT HELIX 196 196
D -> G (in Ref. 2).
S -> G (in Ref. 2).
FG -> VS (in Ref. 2).
VTGV -> TSGS (in Ref. 3).
D -> G (in Ref. 2).
R -> G (in Ref. 2).
N -> D (in Ref. 3).
N -> D (in Ref. 3).
E -> G (in Ref. 2).
ECNLD -> DCDVA (in Ref. 3).
NE -> DQ (in Ref. 3).
E -> G (in Ref. 2).
HYS -> QYA (in Ref. 3).
I -> V (in Ref. 3).
VT -> IV (in Ref. 3).
I -> V (in Ref. 3).
K -> E (in Ref. 2).
H -> Y (in Ref. 2).
P -> S (in Ref. 2).
GVDDGGRYSA -> E (in Ref. 2).

```

```

FT CHAIN 127 346 actinidin.
SQ SEQUENCE 380 AA; 42142 MW; 56322CEE91B72B15 CRC64;

Query Match 84.2%; Score 48; DB 2; Length 380;
Best Local Similarity 77.8%; Pred. No. 5.2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9
DB 304 WIVKNSWDT 312
| | | | |
| | | | |

RESULT 14
Q9AXD2 PRELIMINARY; PRT; 380 AA.
ID Q9AXD2
AC Q9AXD2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Actinidin protease.
OS Actinidia chinensis (Kiwi) (Yangtao).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Ericales; Actinidiaceae; Actinidia.
OC NCBI_TaxID=3625;
RN [1]_
RX SEQUENCE FROM N.A.
RP Lee N.X., Hahn Y.T.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family C1.
EMBL; AF343446; AAK06862.1; -.
DR HSSP; P00785; 1AEC.
DR GO; GO:0004137; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000658; peptidase_C1
DR InterPro; IPR000169; pept_cys_acsite.
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PR00705; PAPAIN.
DR ProDom; PD000158; Peptidase_C1; 1.
DR SMART; SM00645; Pept C1; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; 1.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Hydrolase; Protease; Thiol protease.
SQ SEQUENCE 380 AA; 42172 MW; 26BFD612B53A50B CRC64;

Query Match 84.2%; Score 48; DB 2; Length 380;
Best Local Similarity 77.8%; Pred. No. 5.2;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9
DB 304 WIVKNSWDT 312
| | | | |
| | | | |

RESULT 15
Q96227 PRELIMINARY; PRT; 302 AA.
ID Q96227
AC Q96227;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Actinidin precursor (fragment).
OS Actinidia deliciosa (Kiwi).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Ericales; Actinidiaceae; Actinidia.
OC NCBI_TaxID=3627;
RN [1]_
RX SEQUENCE FROM N.A.
RP Prasekitt U.M., McKee R.A., Smith H.;
RL "Molecular analysis of actinidin, the cysteine proteinase of actinidia
RT chinensis.";
```

RL Plant Mol. Biol. 10:193-202(1988).  
 CC -I- SIMILARITY: Belongs to peptidase family Cl.  
 DR EMBL; M21335; AAA32630.1; -  
 DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000668; Peptidase\_Cl.  
 DR InterPro; IPR00169; Pept\_cys\_acsite.  
 DR Pfam; PF00112; Peptidase\_Cl; 1.  
 DR PRINTS; PR00705; PAPAIN.  
 DR ProDom; PD000158; Peptidase\_Cl; 1.  
 DR SMART; SM00645; Pept\_Cl; 1.  
 DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 KW Hydrolase; Protease; Signal; Thiol protease.  
 FT NON TER 1  
 FT SIGNAL <1 57 Potential.  
 FT CHAIN 58 277 actinidin.  
 SQ SEQUENCE 302 AA; 32745 MW; 99835E2238D620DB CRC64;

Query Match 80.7%; Score 46; DB 2; Length 302;  
 Best Local Similarity 77.8%; Pred. No. 8.9;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 235 WIVNSWDT 243

Search completed: October 27, 2004, 17:50:20  
 Job time : 15.3028 secs

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:26:32 ; Search time 12.4064 Seconds  
(without alignments)  
260.234 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWD 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Genesep23Sep04: \*  
1: Genesep1980s: \*  
2: Genesep1990s: \*  
3: Genesep2000s: \*  
4: Genesep2001s: \*  
5: Genesep2002s: \*  
6: Genesep2003as: \*  
7: Genesep2003bs: \*  
8: Genesep2004s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	9	5	AAC20571
2	57	100.0	9	5	AB988536 Cysteine
3	51	89.5	23	2	AAR36414 DPI-28.2
4	51	89.5	23	2	AAR36476 DPI-28.2
5	51	89.5	23	2	AAR51762 Der p I d
6	51	89.5	23	2	AAR51824 Der f I d
7	51	89.5	23	2	AAR71998 Dermatoph
8	51	89.5	23	2	AAR71937 Dermatoph
9	51	89.5	23	2	AAV50453 Dermatoph
10	51	89.5	23	2	AAV50391 Dermatoph
11	51	89.5	23	4	AAU19056 T-cell ep
12	51	89.5	23	4	AAU18994 T-cell ep
13	51	89.5	24	2	AAR36462 DPI-10(18
14	51	89.5	24	2	AAR36399 DPI-10(18
15	51	89.5	24	2	AAR51747 Der p I d
16	51	89.5	24	2	AAR51810 Der f I d
17	51	89.5	24	2	AAR77134 Dermatoph
18	51	89.5	24	2	AAR71985 Dermatoph
19	51	89.5	24	2	AAR71927 Dermatoph
20	51	89.5	24	2	AAV50439 Dermatoph
21	51	89.5	24	2	AAV50376 Dermatoph
22	51	89.5	24	4	AAU19042 T-cell ep
23	51	89.5	24	4	AAU18979 T-cell ep
24	51	89.5	29	2	AAR36475 DPI-28.1
25	51	89.5	29	2	AAR36413 DPI-28.1

26	51	89.5	29	2	AAR51823 Der f I d
27	51	89.5	29	2	AAR51761 Der p I d
28	51	89.5	29	2	AAW71997 Dermatoph
29	51	89.5	29	2	AAW71936 Dermatoph
30	51	89.5	29	2	AAW50452 Dermatoph
31	51	89.5	29	2	AAV50390 Dermatoph
32	51	89.5	29	4	AAU19055 T-cell ep
33	51	89.5	29	4	AAU18993 T-cell ep
34	51	89.5	72	4	AAU07749 House dus
35	51	89.5	211	2	AAV25677 Euroglyph
36	51	89.5	211	2	AAV25676 Euroglyph
37	51	89.5	211	2	AAV25678 Euroglyph
38	51	89.5	211	7	ADC34925 Euroglyph
39	51	89.5	211	7	ADC34926 Euroglyph
40	51	89.5	212	2	AAV25679 Euroglyph
41	51	89.5	212	7	ADC34927 Euroglyph
42	51	89.5	212	2	AAR52742 Protein a
43	51	89.5	222	2	AAR52742 House dus
44	51	89.5	222	4	AAU07746 House dus
45	51	89.5	222	4	AAU07748 House dus

#### ALIGNMENTS

RESULT 1  
AAC20571  
ID AAC20571 standard; peptide; 9 AA.  
XX AAC20571;  
XX AC  
XX DT 02-JAN-2003 (first entry)  
XX DE Cysteine protease epitope peptide region, SEQ ID No 5.  
XX KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
XX KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
XX KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
XX KW atopic eczema; epitope.  
XX OS Dermatophagoides pteronyssinus.  
XX PN WO200278736-A2.  
XX PD 10-OCT-2002.  
XX PF 28-MAR-2002; 2002WO-FR001098.  
XX PR 30-MAR-2001; 2001FR-00004370.  
XX PR 03-MAY-2001; 2001FR-00005929.  
XX PR 29-MAY-2001; 2001US-00867159.  
XX PA (ANTI-) ANTIALIS SARL.  
XX PI Loria E, Terrasse G, Trehin Y;  
XX WPI; 2002-750636/81.  
XX Antiallergic compositions containing an anti-histamine, a histamine  
XX synthesis inhibitor, and optionally an allergen or nucleic acid coding  
XX for the allergen.  
XX Claim 14; Page 11; 32pp; French.  
XX The invention relates to antiallergic compositions containing an anti-  
XX histamine, a histamine synthesis inhibitor, and optionally an allergen or  
XX isolated nucleic acid molecule that has at least one polynucleotide  
XX sequence coding for the allergen, together with a pharmaceutical carrier.  
XX The pharmaceutical composition of the invention is useful as a non-  
XX specific antiallergic treatment, and also useful in the treatment of  
XX allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
XX allergic and atopic eczema. This sequence represents a peptide of a  
XX cysteine protease epitope region relating to the antiallergic

```

CC compositions of the invention
XX
SQ Sequence 9 AA;

Query Match      100.0%; Score 57; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 1 WTVRNSWDT 9

RESULT 2
ABB98536
ID ABB98536 standard; peptide; 9 AA.
XX
AC ABB98536;
XX
DT 13-DEC-2002 (first entry)
XX
DE Cysteine protease epitope #3.
XX
KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;
KW immunotherapy; allergen; allergic hypersensitivity reaction;
KW allergic asthma; allergic rhinitis; allergic atopic eczema;
KW cysteine protease.
XX
OS Dermatophagoides pteronyssinus.
XX
PN FR2822709-A1.
XX
PD 04-OCT-2002.
XX
PF 03-MAY-2001; 2001FR-00005929.
XX
PR 30-MAR-2001; 2001FR-00004370.
XX
PA (ANTI-) ANTIALIS SARL.
XX
PI Loria E, Terrasse G, Trehin Y;
XX
DR WPI; 2002-735037/80.
XX
PT Antiallergic composition, useful for preventing and treating e.g. asthma,
PT rhinitis or eczema, containing at least two of allergen, antihistamine
PT and histamine synthesis inhibitor.
XX
PS Claim 8; Page 6; 33pp; French.
XX
CC The present invention relates to an antiallergic pharmaceutical
CC composition (i) comprising a pharmaceutical carrier containing an active
CC agent combination of at least two of: an allergen; an antihistamine; and
CC a histamine synthesis inhibitor. (i) is used for treating or preventing
CC allergic hypersensitivity reactions, especially allergic asthma, allergic
CC rhinitis or allergic atopic eczema, in babies, children or adults. The
CC present sequence is a peptide fragment (epitope) of cysteine protease
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the
CC invention
XX
SQ Sequence 9 AA;

Query Match      100.0%; Score 57; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 1 WTVRNSWDT 9

RESULT 3
AAR36414
ID AAR36414 standard; peptide; 23 AA.
XX
AC AAR36414;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
DE DPI-28.2(173-195) a Dermatophagoides protein allergen.
XX
KW T cell epitope; house dust mite; allergy; soluble; Der pl.
XX
OS Synthetic.
XX
PN WO9308279-A1.
XX
PD 29-APR-1993.
XX
PF 15-OCT-1992; 92WO-US008637.
XX
PR 16-OCT-1991; 91US-00777859.
PR 08-MAY-1992; 92US-00881396.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
DR WPI; 1993-152472/18.
XX
PT Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis
PT and treatment of sensitivity to house dust mite.
XX
PS Claim 10; Fig 3; 176pp; English.
XX
CC The peptide is one of a series of overlapping peptides synthesised by
CC standard techniques to cover the whole Dermatophagoides pteronyssinus Der
CC PI sequence. The T cell epitopes of the protein were mapped by detection
CC of the peptide's ability to stimulate T cell activity. The peptides may
CC be used for diagnosis and treatment of sensitivity to house dust mite
CC allergens. When administered to house dust mite sensitive individuals,
CC the peptides are capable of modifying the allergic response to the
CC allergens. The peptides may be modified for e.g. increasing solubility,
CC enhancing therapeutic or preventive efficacy or stability. See also
CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 23 AA;

Query Match      89.5%; Score 51; DB 2; Length 23;
Best Local Similarity 88.9%; Pred. No. 0.24;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 14 WTVRNSWDT 22

RESULT 4
AAR36476
ID AAR36476 standard; peptide; 23 AA.
XX
AC AAR36476;
XX
DT 25-MAR-2003 (revised)
DT 12-AUG-1993 (first entry)
XX
DE DPI-28.2(173-195), a Dermatophagoides protein allergen.
XX
KW T cell epitope; house dust mite; allergy; soluble; Def pl.
XX
OS Synthetic.
XX
PN WO9308279-A1.
XX

```



PD 29-APR-1993.  
 XX 15-OCT-1992; 92WO-US008637.  
 XX 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;  
 XX WPI; 1993-152472/18.  
 XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis  
 PT and treatment of sensitivity to house dust mite.  
 XX Claim 44; Fig 4; 176pp; English.  
 XX The peptide is one of a series of overlapping peptides synthesised by  
 CC standard techniques to cover the whole Dermatophagoides farinae Def PI  
 CC sequence. The T cell epitopes of the protein were mapped by detection of  
 CC the peptide's ability to stimulate T cell activity. The peptides may be  
 CC used for diagnosis and treatment of sensitivity to house dust mite  
 CC allergens. When administered to house dust mite sensitive individuals,  
 CC the peptides are capable of modifying the allergic response to the  
 CC allergens. The peptides may be modified for e.g. increasing solubility,  
 CC enhancing therapeutic or preventive efficacy or stability. See also  
 CC AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX SQ Sequence 23 AA;  
 Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 WTVRNSWDT 9  
 Db 14 WIVRNSWDT 22  
 RESULT 5  
 AAR51762  
 ID AAR51762 standard; protein; 23 AA.  
 XX AAR51762;  
 XX 01-FEB-1995 (first entry)  
 XX Der p I derived peptide, DP I-28.2(173-195).  
 XX Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
 KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;  
 KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
 KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
 KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
 KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
 KW T cell subpopulations; unresponsive; immune response; tolerance.  
 XX Dermatophagoides pteronyssinus.  
 XX ZA9302677-A.  
 XX 26-JAN-1994.  
 XX 16-APR-1993; 93ZA-00002677.  
 XX 16-APR-1993; 93ZA-00002677.  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX PI

DR WPI; 1994-126807/15.  
 XX Isolated and/or modified peptides comprising T-cell epitopes - of major  
 PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
 PT sensitivity to house dust mites.  
 XX Claim 28; Page 67; 154pp; English.  
 XX The sequences given in AAR51731-841 represent T-cell epitopes derived  
 CC from the group I and II protein allergens from the house dust mite D.  
 CC farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II  
 CC respectively. The Der f II proteinsCC shows high homology having an  
 CC identity of 88%, with an identity of 81% between the two group I proteins  
 CC (see also AAR51727-301). Fusion peptides may be produced which comprise at  
 CC least two or these antigenic fragments. Each region of these fusion  
 CC peptides may be derived from the same, or different, mite allergens. The  
 CC antigenic fragments may be altered by substitution, deletion or addition  
 CC to enhance their antigenicity. These peptides may be produced by chemical  
 CC synthesis, chemical cleavage of the protein allergen or by recombinant  
 CC techniques. These peptides, or the fusion peptides, when administered to  
 CC a house dust mite sensitive individual, are capable of modifying the  
 CC allergic response of the individual to the allergen. The peptides do not  
 CC bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the  
 CC full length protein allergen. This reduces the major complications of  
 CC standard immunotherapy, which are IgE-mediated responses such as  
 CC anaphylaxis. Exposure of mite allergic patients to these peptides may  
 CC tolerate or anergise appropriate T cell subpopulations such that they  
 CC become unresponsive to mite allergens and do not participate in mounting  
 CC an immune response upon exposure. Administration of the peptides may also  
 CC modify the lymphokine secretion profile as compared with exposure to the  
 CC naturally occurring mite protein allergen  
 XX SQ Sequence 23 AA;  
 Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 WTVRNSWDT 9  
 Db 14 WIVRNSWDT 22  
 RESULT 6  
 AAR51824  
 ID AAR51824 standard; protein; 23 AA.  
 XX AAR51824;  
 XX 01-FEB-1995 (first entry)  
 XX Der f I derived peptide, DP I-28.2(173-195).  
 XX Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;  
 KW homology; D. farinae; Der f I; group II; Der p II; Der f II; T-cell;  
 KW epitopes; fusion peptides; antigenic fragments; substitution; deletion;  
 KW addition; chemical synthesis; chemical cleavage; recombinant techniques;  
 KW allergic response; immunoglobulin E; IgE; immunotherapy; anaphylaxis;  
 KW IgE-mediated responses; anergise; lymphokine secretion profile; modify;  
 KW T cell subpopulations; unresponsive; immune response; tolerance.  
 XX Dermatophagoides farinae.  
 XX ZA9302677-A.  
 XX 26-JAN-1994.  
 XX 16-APR-1993; 93ZA-00002677.  
 XX 16-APR-1993; 93ZA-00002677.  
 XX (IMMU-) IMMULOGIC PHARM CORP.

PI Garman RD, Greenstein JL, Kuo M, Rogers BL;  
 XX WPI; 1994-126807/15.  
 XX  
 XX Isolated and/or modified peptides comprising T-cell epitopes - of major  
 PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
 PT sensitivity to house dust mites.  
 XX  
 XX Claim 28; Page 86; 154pp; English.  
 XX  
 XX The sequences given in AAR51731-841 represent T-cell epitopes derived  
 CC from the group I and II protein allergens from the house dust mite D.  
 CC farinae and D. pteronyssinus. Der f I, Der f II, Der p I and Der p II  
 CC respectively. The Der f II proteinsCC shows high homology having an  
 CC identity of 88%, with an identity of 81% between the two group I proteins  
 CC (see also AAR51727-30). Fusion peptides may be produced which comprise at  
 CC least two or these antigenic fragments. Each region of these fusion  
 CC peptides may be derived from the same, or different, mite allergens. The  
 CC antigenic fragments may be altered by substitution, deletion or addition  
 CC to enhance their antigenicity. These peptides may be produced by chemical  
 CC synthesis, chemical cleavage of the protein allergen or by recombinant  
 CC techniques. These peptides, or the fusion peptides, when administered to  
 CC a house dust mite sensitive individual, are capable of modifying the  
 CC allergic response of the individual to the allergen. The peptides do not  
 CC bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the  
 CC full length protein allergen. This reduces the major complications of  
 CC standard immunotherapy, which are IgE-mediated responses such as  
 CC anaphylaxis. Exposure of mite allergic patients to these peptides may  
 CC tolerate or anergise appropriate T cell subpopulations such that they  
 CC become unresponsive to mite allergens and do not participate in mounting  
 CC an immune response upon exposure. Administration of the peptides may also  
 CC modify the lymphokine secretion profile as compared with exposure to the  
 CC naturally occurring mite protein allergen  
 XX  
 XX Sequence 23 AA;

Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9  
 Db 14 WIVRNSWDT 22

RESULT 7  
 AAW71998  
 ID AAW71998 standard; peptide; 23 AA.

XX AAW71998;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 16-DEC-1998 (first entry)  
 XX  
 XX Dermatophagoides Der f I protein peptide DFI-28.2.  
 DE  
 XX genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
 KW Der p II; Der f I; Der f II; house dust mite allergy.  
 KW  
 XX Dermatophagoides.  
 OS  
 XX US5820862-A.  
 PN  
 XX 13-OCT-1998.  
 PD  
 XX 07-JUN-1995; 95US-00482142.  
 PF  
 XX 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 PR  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
 PI Shaked Z, Rogers BL;  
 PI  
 XX WPI; 1998-567590/48.  
 DR  
 XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.  
 PT  
 XX Disclosure; Col 97-99; 155pp; English.  
 PS  
 XX The present invention describes peptides for treating sensitivity to

XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
 PI Shaked Z, Rogers BL;  
 XX  
 XX WPI; 1998-567590/48.  
 XX  
 XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.  
 PT  
 XX Disclosure; Col 129-130; 155pp; English.  
 PS  
 XX The present invention describes peptides for treating sensitivity to  
 CC house dust mite allergens from the genus Dermatophagoides. Peptides  
 CC within the scope of the invention comprise at least one T cell epitope,  
 CC or preferably at least two T cell epitopes of a protein allergen selected  
 CC from the allergens Der p I, Der p II, Der f I, or Der f II. The invention  
 CC also describes modified peptides having similar or enhanced therapeutic  
 CC properties as the corresponding, naturally occurring allergen, but having  
 CC reduced side effects. AAW71912 to AAW72000, and AAW72257 to AAW72330  
 CC represent peptides from the present invention. (Updated on 25-MAR-2003 to  
 CC correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 23 AA;

Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9  
 Db 14 WIVRNSWDT 22

RESULT 8  
 AAW71937  
 ID AAW71937 standard; peptide; 23 AA.

XX AAW71937;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 16-DEC-1998 (first entry)  
 XX  
 XX Dermatophagoides Der p I protein peptide DFI-28.2.  
 DE  
 XX genus Dermatophagoides; major protein allergen; T cell epitope; Der p I;  
 KW Der p II; Der f I; Der f II; house dust mite allergy.  
 KW  
 XX Dermatophagoides.  
 OS  
 XX US5820862-A.  
 PN  
 XX 13-OCT-1998.  
 PD  
 XX 07-JUN-1995; 95US-00482142.  
 PF  
 XX 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 PR  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 PA  
 XX Franzen HM, Kuo M, Evans S, Garman RD, Greenstein JL, Chen X;  
 PI Shaked Z, Rogers BL;  
 PI  
 XX WPI; 1998-567590/48.  
 DR  
 XX Dermatophagoides allergen peptides - useful for treating house dust mite  
 PT allergy.  
 PT  
 XX Disclosure; Col 97-99; 155pp; English.  
 PS  
 XX The present invention describes peptides for treating sensitivity to

CC house dust mite allergens from the genus Dermatophagoides. Peptides  
CC within the scope of the invention comprise at least one T cell epitope,  
CC or preferably at least two T cell epitopes of a protein allergen selected  
CC from the allergens Der p I, Der p II, Der f I, or Der f II. The invention  
CC also describes modified peptides having similar or enhanced therapeutic  
CC properties as the corresponding, naturally occurring allergen, but having  
CC reduced side effects. AAW71912 to AAW72000, and AAW72257 to AAW72330  
CC represent peptides from the present invention. (Updated on 25-MAR-2003 to  
CC correct PR field.) (Updated on 27-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 23 AA;  
Query Match 89.5%; Score 51; DB 2; Length 23;  
Best Local Similarity 88.9%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WTVNSWDT 9  
DB 14 WIVNSWDT 22  
RESULT 9  
AAAY50453  
ID AAY50453 standard; peptide; 23 AA.  
XX  
AC AAY50453;  
XX  
DT 25-JAN-2000 (first entry)  
XX  
DE Dermatophagoides sp major protein allergen DP I-28.2.  
XX  
KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
KW Der f I; Der p I; Der p II; Der f II.  
XX  
OS Dermatophagoides sp.  
XX  
FN US968526-A.  
XX  
PD 19-OCT-1999.  
XX  
PF 07-JUN-1995; 95US-00478572.  
XX  
PR 14-APR-1994; 94US-00227772.  
PR 12-APR-1995; 95WO-US004481.  
PR 19-MAY-1995; 95US-00445307.  
XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX  
PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;  
PI Chen X, Evans S, Kuo M;  
XX  
DR WPI; 1999-590385/50.  
XX  
PT Screening individuals for allergic reactions to T cell epitopes of major  
XX allergens from house dust mites.  
XX  
PS Claim 3m; Col 129-130; 158pp; English.  
XX  
CC This invention describes a novel method (I) for detecting whether an  
CC individual is sensitive to Dermatophagoides (house dust mites). The  
CC method involves detecting sensitivity to house dust mites in patients,  
CC comprising combining a blood sample from the individual with 1 or more  
CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
CC with varying, defined amino acids sequences (given in the specification)  
CC may be used in (I). The sample and allergens are combined under  
CC conditions appropriate for the binding of blood components with the  
CC polypeptides. The extent of binding is then indicative of the sensitivity  
CC of the patient to house dust mites. (I) may be used to screen individuals  
CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
CC mite is a major cause of a variety of allergic disorders such as asthma,  
CC rhinitis and ectopic dermatitis. AAY50360-Y50542 and AAY50546-Y50555

CC represent house dust mite allergen peptide fragments derived from Der p  
CC I, Der f II, Der f I and Der f II  
XX  
SQ Sequence 23 AA;  
Query Match 89.5%; Score 51; DB 2; Length 23;  
Best Local Similarity 88.9%; Pred. No. 0.24;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WTVNSWDT 9  
DB 14 WIVNSWDT 22  
RESULT 10  
AAAY50391  
ID AAY50391 standard; peptide; 23 AA.  
XX  
AC AAY50391;  
XX  
DT 25-JAN-2000 (first entry)  
XX  
DE Dermatophagoides sp major protein allergen DP I-28.2.  
XX  
KW Allergen; house dust mite; detection; sensitivity; T cell epitope;  
KW screening; allergic disorder; asthma; rhinitis; ectopic dermatitis;  
KW Der f I; Der p I; Der p II; Der f II.  
XX  
OS Dermatophagoides sp.  
XX  
FN US968526-A.  
XX  
PD 19-OCT-1999.  
XX  
PF 07-JUN-1995; 95US-00478572.  
XX  
PR 14-APR-1994; 94US-00227772.  
PR 12-APR-1995; 95WO-US004481.  
PR 19-MAY-1995; 95US-00445307.  
XX  
PA (IMMU-) IMMULOGIC PHARM CORP.  
XX  
PI Garman RD, Greenstein JL, Rogers BL, Franzen HM, Shaked Z;  
PI Chen X, Evans S, Kuo M;  
XX  
DR WPI; 1999-590385/50.  
XX  
PT Screening individuals for allergic reactions to T cell epitopes of major  
XX allergens from house dust mites.  
XX  
PS Claim 1e'; Column 99-100; 158pp; English.  
XX  
CC This invention describes a novel method (I) for detecting whether an  
CC individual is sensitive to Dermatophagoides (house dust mites). The  
CC method involves detecting sensitivity to house dust mites in patients,  
CC comprising combining a blood sample from the individual with 1 or more  
CC isolated T cell epitopes of the protein allergens I and II ((DP I) and  
CC (DP II)) from Dermatophagoides (house dust mites). 32 T cell epitopes  
CC with varying, defined amino acids sequences (given in the specification)  
CC may be used in (I). The sample and allergens are combined under  
CC conditions appropriate for the binding of blood components with the  
CC polypeptides. The extent of binding is then indicative of the sensitivity  
CC of the patient to house dust mites. (I) may be used to screen individuals  
CC for sensitivity to Dermatophagoides (house dust mites). The house dust  
CC mite is a major cause of a variety of allergic disorders such as asthma,  
CC rhinitis and ectopic dermatitis. AAY50360-Y50542 and AAY50546-Y50555  
XX  
SQ Sequence 23 AA;  
Query Match 89.5%; Score 51; DB 2; Length 23;  
Best Local Similarity 88.9%; Pred. No. 0.24;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 14 WIVNSWDT 22

RESULT 11  
 AAU19056  
 ID AAU19056 standard; peptide; 23 AA.  
 AC AAU19056;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE T-cell epitope containing peptide DFI-28.2.  
 XX  
 KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX  
 OS Dermatophagoides farinae.  
 XX  
 PN US6268491-B1.  
 XX  
 PD 31-JUL-2001.  
 XX  
 PF 07-JUN-1995; 95US-00484296.  
 XX  
 PR 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 PI  
 DR WPI; 2001-549074/61.  
 XX  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX  
 PS Claim 1; Col 129; 158pp; English.  
 XX  
 CC The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic T-cell epitope containing peptide  
 CC derived from the Dermatophagoides allergenic proteins  
 XX  
 SQ Sequence 23 AA;

Query Match 89.5%; Score 51; DB 4; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 14 WIVNSWDT 22

RESULT 12  
 AAU18994  
 ID AAU18994 standard; peptide; 23 AA.  
 AC AAU18994;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE T-cell epitope containing peptide DPI-28.2.  
 XX  
 KW House dust mite; allergenic peptide; Der p I; Der p II; Der f I;  
 KW Der f II; antiallergenic; immunostimulant; house dust mite allergy;  
 KW T-cell epitope.  
 XX  
 OS Dermatophagoides pteronyssinus.  
 XX  
 PN US6268491-B1.  
 XX  
 PD 31-JUL-2001.  
 XX  
 PF 07-JUN-1995; 95US-00484296.  
 XX  
 PR 16-OCT-1991; 91US-00777859.  
 PR 08-MAY-1992; 92US-00881396.  
 PR 14-APR-1993; 93WO-US003471.  
 PR 14-APR-1994; 94US-00227772.  
 PR 19-MAY-1995; 95US-00445307.  
 XX  
 XX (IMMU-) IMMULOGIC PHARM CORP.  
 XX  
 PI Garman RD, Greenstein JL, Kuo M, Rogers BL, Franzen HM, Chen X;  
 PI Evans S, Shaked Z;  
 PI  
 DR WPI; 2001-549074/61.  
 XX  
 XX Peptides comprising T cell groups of the major allergens from  
 PT Dermatophagoides (house dust mites), useful for treating house dust mite  
 PT allergy in humans, and for diagnosing sensitivity to house dust mite  
 PT protein allergens.  
 XX  
 PS Claim 1; Fig 3; 158pp; English.  
 XX  
 CC The invention relates to an isolated peptide of the major protein  
 CC allergens of the genus Dermatophagoides, which comprises at least one T  
 CC cell group of a protein allergen from Der p (DP) I, DP II, Der f (DF) I  
 CC or DF II. The isolated peptide comprises at least two regions, each  
 CC region comprising at least one T cell group of a protein allergen of the  
 CC genus Dermatophagoides. The regions are derived from the same or  
 CC different protein allergens of the genus Dermatophagoides. The peptides  
 CC are useful for treating house dust mite allergy in humans. The peptides  
 CC are also useful for detecting or diagnosing sensitivity to house dust  
 CC mite protein allergens. The present peptides have similar or enhanced  
 CC therapeutic properties as the naturally-occurring allergen, but have  
 CC reduced side effects, and increased solubility and stability. The present  
 CC sequence represents an allergenic T-cell epitope containing peptide  
 CC derived from the Dermatophagoides allergenic proteins  
 XX  
 SQ Sequence 23 AA;

Query Match 89.5%; Score 51; DB 4; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.24;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9  
 Db 14 WIVNSWDT 22

RESULT 13  
 AAR36462  
 ID AAR36462 standard; peptide; 24 AA.

```

XX AAR36462;
XX
XX 25-MAR-2003 (revised)
XX
XX 12-AUG-1993 (first entry)
XX
XX DFI-10(181-204), a Dermatophagoides protein allergen.
XX
XX T cell epitope; house dust mite; allergy; soluble; Def pI.
XX
XX Synthetic.
XX
XX WO9308279-A1.
XX
XX 29-APR-1993.
XX
XX 15-OCT-1992; 92WO-US008637.
XX
XX 16-OCT-1991; 91US-00777859.
XX
XX 08-MAY-1992; 92US-00881396.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
XX WPI; 1993-152472/18.
XX
XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis
XX and treatment of sensitivity to house dust mite.
XX
XX Claim 44; Fig 4; 176pp; English.
XX
XX The peptide is one of a series of overlapping peptides synthesised by
XX standard techniques to cover the whole Dermatophagoides farinae Def pI
XX sequence. The T cell epitopes of the protein were mapped by detection of
XX the peptide's ability to stimulate T cell activity. The peptides may be
XX used for diagnosis and treatment of sensitivity to house dust mite
XX allergens. When administered to house dust mite sensitive individuals,
XX the peptides are capable of modifying the allergic response to the
XX allergens. The peptides may be modified for e.g. increasing solubility,
XX enhancing therapeutic or preventive efficacy or stability. See also
XX AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
XX Sequence 24 AA;
XX
XX Query Match 89.5%; Score 51; DB 2; Length 24;
XX Best Local Similarity 88.9%; Pred. No. 0.25;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 WTVRNSWDT 9
XX | |||||
XX 6 WIVRNSWDT 14
XX
XX RESULT 15
XX AAR51747
XX ID AAR51747 standard; peptide; 24 AA.
XX
XX AC AAR51747;
XX
XX 01-FEB-1995 (first entry)
XX
XX Der p I derived peptide, DP I-10(181-204).
XX
XX Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;
XX homology; D. farinae; Der f; group II; Der p II; Der f II; T-cell;
XX epitopes; fusion peptides; antigenic fragments; substitution; deletion;
XX addition; chemical synthesis; chemical cleavage; recombinant techniques;
XX allergic response; immunoglobulin E; IGE; immunotherapy; anaphylaxis;
XX IGE-mediated responses; anergise; lymphokine secretion profile; modify;
XX T cell subpopulations; unresponsive; immune response; tolerance.
XX
XX Dermatophagoides pteronyssinus.
XX
XX ZA9302677-A.
XX
XX 26-JAN-1994.
XX
XX 16-APR-1993; 93ZA-00002677.
XX
XX 16-APR-1993; 93ZA-00002677.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo M, Rogers BL;
XX
XX WPI; 1994-126807/15.

```

```

XX PF 15-OCT-1992; 92WO-US008637.
XX
XX 16-OCT-1991; 91US-00777859.
XX
XX 08-MAY-1992; 92US-00881396.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo MC, Rogers BL;
XX
XX WPI; 1993-152472/18.
XX
XX Isolated peptide(s) of dermatophagoides protein allergens - for diagnosis
XX and treatment of sensitivity to house dust mite.
XX
XX Claim 10; Fig 3; 176pp; English.
XX
XX The peptide is one of a series of overlapping peptides synthesised by
XX standard techniques to cover the whole Dermatophagoides pteronyssinus Der
XX pI sequence. The T cell epitopes of the protein were mapped by detection
XX of the peptide's ability to stimulate T cell activity. The peptides may
XX be used for diagnosis and treatment of sensitivity to house dust mite
XX allergens. When administered to house dust mite sensitive individuals,
XX the peptides are capable of modifying the allergic response to the
XX allergens. The peptides may be modified for e.g. increasing solubility,
XX enhancing therapeutic or preventive efficacy or stability. See also
XX AAR34686-700 and AAR36398-490. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
XX Sequence 24 AA;
XX
XX Query Match 89.5%; Score 51; DB 2; Length 24;
XX Best Local Similarity 88.9%; Pred. No. 0.25;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 WTVRNSWDT 9
XX | |||||
XX 6 WIVRNSWDT 14
XX
XX RESULT 15
XX AAR51747
XX ID AAR51747 standard; protein; 24 AA.
XX
XX AC AAR51747;
XX
XX 01-FEB-1995 (first entry)
XX
XX Der p I derived peptide, DP I-10(181-204).
XX
XX Group I; protein allergen; house dust mite; D. pteronyssinus; Der p I;
XX homology; D. farinae; Der f; group II; Der p II; Der f II; T-cell;
XX epitopes; fusion peptides; antigenic fragments; substitution; deletion;
XX addition; chemical synthesis; chemical cleavage; recombinant techniques;
XX allergic response; immunoglobulin E; IGE; immunotherapy; anaphylaxis;
XX IGE-mediated responses; anergise; lymphokine secretion profile; modify;
XX T cell subpopulations; unresponsive; immune response; tolerance.
XX
XX Dermatophagoides pteronyssinus.
XX
XX ZA9302677-A.
XX
XX 26-JAN-1994.
XX
XX 16-APR-1993; 93ZA-00002677.
XX
XX 16-APR-1993; 93ZA-00002677.
XX
XX (IMMU-) IMMULOGIC PHARM CORP.
XX
XX Garman RD, Greenstein JL, Kuo M, Rogers BL;
XX
XX WPI; 1994-126807/15.

```

XX Isolated and/or modified peptides comprising T-cell epitopes - of major  
PT protein allergens of genus Dermatophagoides, used to treat or diagnose  
PT sensitivity to house dust mites.  
XX  
PS Disclosure; Fig 3; 154pp; English.  
XX  
CC The sequences given in AAR51731-841 represent T-cell epitopes derived  
CC from the group I and II protein allergens from the house dust mite D.  
CC farinae and D. pteronyssinus, Der f I, Der f II, Der p I and Der p II  
CC respectively. The Der f II proteinsCC shows high homology having an  
CC identity of 88%, with an identity of 81% between the two group I proteins  
CC (see also AAR51727-30). Fusion peptides may be produced which comprise at  
CC least two or these antigenic fragments. Each region of these fusion  
CC peptides may be derived from the same, or different, mite allergens. The  
CC antigenic fragments may be altered by substitution, deletion or addition  
CC to enhance their antigenicity. These peptides may be produced by chemical  
CC synthesis, chemical cleavage of the protein allergen or by recombinant  
CC techniques. These peptides, or the fusion peptides, when administered to  
CC a house dust mite sensitive individual, are capable of modifying the  
CC allergic response of the individual to the allergen. The peptides do not  
CC bind to immunoglobulin E (IgE), or bind IgE to a lesser extent than the  
CC full length protein allergen. This reduces the major complications of  
CC standard immunotherapy, which are IgE-mediated responses such as  
CC anaphylaxis. Exposure of mite allergic patients to these peptides may  
CC tolerate or anergise appropriate T cell subpopulations such that they  
CC become unresponsive to mite allergens and do not participate in mounting  
CC an immune response upon exposure. Administration of the peptides may also  
CC modify the lymphokine secretion profile as compared with exposure to the  
CC naturally occurring mite protein allergen  
XX

SQ Sequence 24 AA;

Query Match 89.5%; Score 51; DB 2; Length 24;  
Best Local Similarity 88.9%; Pred.No. 0.25;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 WTVRNSWDT 9  
| | | | |  
Db 6 WIVRNSWDT 14

Search completed: October 27, 2004, 17:43:58  
Job time : 13.4064 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:50:34 ; Search time 9.32271 Seconds  
(without alignments)  
312.993 Million cell updates/sec

Title: US-09-867-159A-5

Perfect score: 57

Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep.\*  
19: /cgn2\_6/prodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
20: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	9	10	US-09-867-159A-5
2	51	89.5	210	14	Sequence 5, Appl
3	51	89.5	211	10	Sequence 185, App
4	51	89.5	211	14	Sequence 95, Appl
5	51	89.5	211	14	Sequence 184, App
6	51	89.5	222	10	Sequence 2, Appl
7	51	89.5	222	14	Sequence 14, Appl
8	51	89.5	222	14	Sequence 16, Appl
9	51	89.5	222	14	Sequence 18, Appl
10	51	89.5	222	14	Sequence 20, Appl
11	51	89.5	222	14	Sequence 22, Appl
12	51	89.5	222	14	Sequence 24, Appl
13	51	89.5	222	14	Sequence 28, Appl

14	51	89.5	222	14	US-10-001-245-30	Sequence 30, Appl
15	51	89.5	222	14	US-10-001-245-88	Sequence 88, Appl
16	51	89.5	246	14	US-10-001-245-181	Sequence 181, Appl
17	51	89.5	320	9	US-09-877-160-2	Sequence 2, Appl
18	51	89.5	320	10	US-09-847-208-79	Sequence 79, Appl
19	51	89.5	320	14	US-10-001-245-179	Sequence 179, Appl
20	51	89.5	321	10	US-09-847-208-73	Sequence 73, Appl
21	51	89.5	321	14	US-10-001-245-180	Sequence 180, App
22	51	89.5	321	14	US-10-001-245-183	Sequence 183, App
23	51	89.5	327	14	US-10-001-245-182	Sequence 182, App
24	46	80.7	379	14	US-10-369-493-6802	Sequence 6802, Ap
25	46	80.7	657	15	US-10-634-548-49	Sequence 49, Appl
26	46	80.7	657	16	US-10-437-963-179783	Sequence 179783,
27	46	80.7	803	15	US-10-634-548-48	Sequence 48, Appl
28	46	80.7	803	16	US-10-437-963-179784	Sequence 179784,
29	44	77.2	249	15	US-10-425-114-44401	Sequence 44401, A
30	44	77.2	301	15	US-10-424-599-201575	Sequence 201575,
31	44	77.2	381	15	US-10-425-114-43777	Sequence 43777, A
32	43	75.4	361	15	US-10-424-599-270756	Sequence 270756,
33	43	75.4	366	16	US-10-437-963-142915	Sequence 142915,
34	42	73.7	341	15	US-10-424-599-144685	Sequence 144685,
35	42	73.7	354	15	US-10-425-114-40355	Sequence 40355, A
36	42	73.7	362	16	US-10-437-963-162454	Sequence 162454,
37	41	71.9	122	16	US-10-767-701-48766	Sequence 48766, A
38	41	71.9	124	15	US-10-425-114-52357	Sequence 52357, A
39	41	71.9	142	15	US-10-424-599-183258	Sequence 183258,
40	41	71.9	205	15	US-10-425-114-53588	Sequence 53588, A
41	41	71.9	209	16	US-10-437-963-195558	Sequence 195558,
42	41	71.9	218	11	US-09-972-211-90	Sequence 90, Appl
43	41	71.9	218	15	US-10-096-625-90	Sequence 90, Appl
44	41	71.9	220	11	US-09-972-211-89	Sequence 89, Appl
45	41	71.9	220	15	US-10-096-625-89	Sequence 89, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-867-159A-5  
; Sequence 5, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(9)  
; OTHER INFORMATION: Comprises epitope from cystine protease.

Query Match 100.0%; Score 57; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTVRNSWDT 9

Db 1 WTVRNSWDT 9

RESULT 2  
 US-10-001-245-185  
 ; Sequence 185, Application US/10001245  
 ; Publication No. US20030175312A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLM, Jens  
 ; APPLICANT: IPSEN, Henrik  
 ; APPLICANT: LARSEN, Jorgen N.  
 ; APPLICANT: SPANGFORT, Michael D.  
 ; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
 ; FILE REFERENCE: 4305/1H942-US2  
 ; CURRENT APPLICATION NUMBER: US/10/001,245  
 ; PRIOR FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/298,170  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/249,361  
 ; PRIOR FILING DATE: 2000-11-16  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 185  
 ; LENGTH: 210  
 ; TYPE: PRT  
 ; ORGANISM: Dermatophagoides farinae  
 US-10-001-245-185

Query Match 89.5%; Score 51; DB 14; Length 210;  
 Best Local Similarity 88.9%; Pred. No. 3.9;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 186 WIVRNSWDT 194

RESULT 3  
 US-09-847-208-95  
 ; Sequence 95, Application US/09847208  
 ; Publication No. US20030082190A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saxon, Andrew  
 ; APPLICANT: Zhang, Ke  
 ; APPLICANT: Zhu, Daocheng  
 ; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF  
 ; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES  
 ; FILE REFERENCE: UC67.002A  
 ; CURRENT APPLICATION NUMBER: US/09/847,208  
 ; CURRENT FILING DATE: 2001-05-01  
 ; NUMBER OF SEQ ID NOS: 177  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 95  
 ; LENGTH: 211  
 ; TYPE: PRT  
 ; ORGANISM: Euroglyphus maynei (House-dust mite)  
 US-09-847-208-95

Query Match 89.5%; Score 51; DB 10; Length 211;  
 Best Local Similarity 88.9%; Pred. No. 4;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 187 WIVRNSWDT 195

RESULT 4  
 US-10-001-245-184  
 ; Sequence 184, Application US/10001245  
 ; Publication No. US20030175312A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLM, Jens  
 ; APPLICANT: IPSEN, Henrik  
 ; APPLICANT: LARSEN, Jorgen N.  
 ; APPLICANT: SPANGFORT, Michael D.

; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
 ; FILE REFERENCE: 4305/1H942-US2  
 ; CURRENT APPLICATION NUMBER: US/10/001,245  
 ; CURRENT FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/298,170  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/249,361  
 ; PRIOR FILING DATE: 2000-11-16  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 184  
 ; LENGTH: 211  
 ; TYPE: PRT  
 ; ORGANISM: Euroglyphus maynei  
 US-10-001-245-184

Query Match 89.5%; Score 51; DB 14; Length 211;  
 Best Local Similarity 88.9%; Pred. No. 4;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 187 WIVRNSWDT 195

RESULT 5  
 US-09-867-159A-2  
 ; Sequence 2, Application US/09867159A  
 ; Publication No. US20030104013A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANTIALIS  
 ; APPLICANT: TERRASSE, GAETAN  
 ; APPLICANT: LORIA, EMILE  
 ; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
 ; TITLE OF INVENTION: and at least one anti-histamine compound  
 ; FILE REFERENCE: B112812US-antiallis  
 ; CURRENT APPLICATION NUMBER: US/09/867,159A  
 ; CURRENT FILING DATE: 2001-05-29  
 ; PRIOR APPLICATION NUMBER: FR01/04370  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: FR01/05929  
 ; PRIOR FILING DATE: 2001-05-03  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 222  
 ; TYPE: PRT  
 ; ORGANISM: Dermatophagoides pteronyssinus  
 ; FEATURE:  
 ; NAME/KEY: PEPTIDE  
 ; LOCATION: (1)..(222)  
 ; OTHER INFORMATION: Peptide sequence from cystine protease.  
 US-09-867-159A-2

Query Match 89.5%; Score 51; DB 10; Length 222;  
 Best Local Similarity 88.9%; Pred. No. 4.2;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 186 WIVRNSWDT 194

RESULT 6  
 US-10-001-245-14  
 ; Sequence 14, Application US/10001245  
 ; Publication No. US20030175312A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLM, Jens  
 ; APPLICANT: IPSEN, Henrik  
 ; APPLICANT: LARSEN, Jorgen N.  
 ; APPLICANT: SPANGFORT, Michael D.  
 ; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
 ; FILE REFERENCE: 4305/1H942-US2  
 ; CURRENT APPLICATION NUMBER: US/10/001,245



```
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-14

Query Match      89.5%; Score 51; DB 14; Length 222;
Best Local Similarity 88.9%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 186 WIVRNSWDT 194

RESULT 7
US-10-001-245-16
; Sequence 16, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-16

Query Match      89.5%; Score 51; DB 14; Length 222;
Best Local Similarity 88.9%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 186 WIVRNSWDT 194

RESULT 8
US-10-001-245-18
; Sequence 18, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-18

Query Match      89.5%; Score 51; DB 14; Length 222;
Best Local Similarity 88.9%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 186 WIVRNSWDT 194

RESULT 9
US-10-001-245-20
; Sequence 20, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-20

Query Match      89.5%; Score 51; DB 14; Length 222;
Best Local Similarity 88.9%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WTVRNSWDT 9
Db 186 WIVRNSWDT 194

RESULT 10
US-10-001-245-22
; Sequence 22, Application US/10001245
; Publication No. US20030175312A1
; GENERAL INFORMATION:
; APPLICANT: HOLM, Jens
; APPLICANT: IPSEN, Henrik
; APPLICANT: LARSEN, Jorgen N.
; APPLICANT: SPANGFORT, Michael D.
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens
; FILE REFERENCE: 4305/1H942-US2
; CURRENT APPLICATION NUMBER: US/10/001,245
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/298,170
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/249,361
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-10-001-245-22
```

; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-22

Query Match 89.5%; Score 51; DB 14; Length 222;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
| | | | | | | | | |  
Db 186 WIVRNSWDT 194

## RESULT 11

US-10-001-245-24  
; Sequence 24, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/LH942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 24  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-24

Query Match 89.5%; Score 51; DB 14; Length 222;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
| | | | | | | | | |  
Db 186 WIVRNSWDT 194

## RESULT 12

US-10-001-245-26  
; Sequence 26, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/LH942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 26  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-26

Query Match 89.5%; Score 51; DB 14; Length 222;  
Best Local Similarity 88.9%; Pred. No. 4.2;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
| | | | | | | | | |  
Db 186 WIVRNSWDT 194

## RESULT 13

US-10-001-245-28  
; Sequence 28, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/LH942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 28  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-28

Query Match 89.5%; Score 51; DB 14; Length 222;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
| | | | | | | | | |  
Db 186 WIVRNSWDT 194

## RESULT 14

US-10-001-245-30  
; Sequence 30, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/LH942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 30  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-30

Query Match 89.5%; Score 51; DB 14; Length 222;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
| | | | | | | | | |  
Db 186 WIVRNSWDT 194

RESULT 15  
US-10-001-245-88  
; Sequence 88, Application US/10001245  
; Publication No. US20030175312A1  
; GENERAL INFORMATION:  
; APPLICANT: HOLM, Jens  
; APPLICANT: IPSEN, Henrik  
; APPLICANT: LARSEN, Jorgen N.  
; APPLICANT: SPANGFORT, Michael D.  
; TITLE OF INVENTION: No. US20030175312A1 mutant allergens  
; FILE REFERENCE: 4305/IH942-US2  
; CURRENT APPLICATION NUMBER: US/10/001,245  
; CURRENT FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/298,170  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/249,361  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 88  
; LENGTH: 222  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
US-10-001-245-88

Query Match 89.5%; Score 51; DB 14; Length 222;  
Best Local Similarity 88.9%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WTVRNSWDT 9  
Db 186 WIVRNSWDT 194

Search completed: October 27, 2004, 18:15:55  
Job time : 20.3227 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 17:35:57 ; Search time 3.19124 Seconds  
(without alignments)  
187.032 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents:AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	89.5	23	2	US-08-482-142-40
2	51	89.5	23	2	US-08-482-142-102
3	51	89.5	23	2	US-08-478-572-40
4	51	89.5	23	2	US-08-478-572-102
5	51	89.5	23	3	US-08-484-296-40
6	51	89.5	23	3	US-08-484-296-102
7	51	89.5	24	2	US-08-482-142-25
8	51	89.5	24	2	US-08-482-142-88
9	51	89.5	24	2	US-08-478-572-25
10	51	89.5	24	2	US-08-478-572-88
11	51	89.5	24	3	US-08-484-296-25
12	51	89.5	24	3	US-08-484-296-88
13	51	89.5	24	5	PCT-US95-04481-16
14	51	89.5	29	2	US-08-482-142-39
15	51	89.5	29	2	US-08-482-142-101
16	51	89.5	29	2	US-08-478-572-39
17	51	89.5	29	2	US-08-478-572-101
18	51	89.5	29	3	US-08-484-296-39
19	51	89.5	29	3	US-08-484-296-101
20	51	89.5	222	1	US-07-945-288-11
21	51	89.5	222	1	US-08-462-831-11
22	51	89.5	222	1	US-08-461-809-11
23	51	89.5	222	1	US-08-461-441-11
24	51	89.5	222	5	PCT-US93-08518-11
25	51	89.5	245	1	US-07-945-288-2
26	51	89.5	245	1	US-08-462-831-2
27	51	89.5	245	1	US-08-461-809-2

28	51	89.5	245	1	US-08-461-441-2	Sequence 2, Appli
29	51	89.5	245	2	US-08-482-142-2	Sequence 2, Appli
30	51	89.5	245	2	US-08-478-572-2	Sequence 2, Appli
31	51	89.5	245	3	US-08-460-040-2	Sequence 2, Appli
32	51	89.5	245	3	US-08-484-296-2	Sequence 2, Appli
33	51	89.5	245	5	PCT-US93-08518-2	Sequence 2, Appli
34	51	89.5	320	1	US-07-945-288-10	Sequence 10, Appl
35	51	89.5	320	1	US-08-462-831-10	Sequence 10, Appl
36	51	89.5	320	1	US-08-461-809-10	Sequence 10, Appl
37	51	89.5	320	1	US-08-461-441-10	Sequence 10, Appl
38	51	89.5	320	5	PCT-US93-08518-10	Sequence 10, Appl
39	51	89.5	321	1	US-07-945-288-6	Sequence 6, Appli
40	51	89.5	321	1	US-08-462-831-6	Sequence 6, Appli
41	51	89.5	321	1	US-08-461-809-6	Sequence 6, Appli
42	51	89.5	321	1	US-08-461-441-6	Sequence 6, Appli
43	51	89.5	321	2	US-08-482-142-6	Sequence 6, Appli
44	51	89.5	321	2	US-08-478-572-6	Sequence 6, Appli
45	51	89.5	321	3	US-08-484-296-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-08-482-142-40  
; Sequence 40, Application US/08482142  
; Patent No. 5820862  
; GENERAL INFORMATION:  
; APPLICANT: Garman, Richard  
; APPLICANT: Greenstein, Julia  
; APPLICANT: Kuo, Mei-Chang  
; APPLICANT: Rogers, Bruce  
; APPLICANT: Franzen, Henry  
; APPLICANT: Chen, Xian  
; APPLICANT: Evans, Sean  
; APPLICANT: Shaked, Ze'ev  
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
; NUMBER OF SEQUENCES: 207  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: 610 LINCOLN STREET  
; CITY: WALTHAM  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,142  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,307  
; FILING DATE: 07 June 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG, ANNE I.  
; REFERENCE/DOCKET NUMBER: 32,976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 466-6000  
; TELEFAX: (617) 466-6040  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
US-08-482-142-40

Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.11;  
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 14 WIVRNSWDT 22

RESULT 2  
 US-08-482-142-102  
 ; Sequence 102, Application US/08482142  
 ; Patent No. 5820862  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/482,142  
 ; FILING DATE: 07-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION NUMBER:  
 ; FILING DATE: 07 June 1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 102:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-482-142-102

Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.11;  
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 14 WIVRNSWDT 22

RESULT 3  
 US-08-478-572-40  
 ; Sequence 40, Application US/08478572

Patent No. 5968526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 ; STREET: 610 LINCOLN STREET  
 ; CITY: WALTHAM  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII TEXT  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/478,572  
 ; FILING DATE: 07-June-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/445,307  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CRAIG, ANNE I.  
 ; REGISTRATION NUMBER: 32,976  
 ; REFERENCE/DOCKET NUMBER: 017.6US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 466-6000  
 ; TELEFAX: (617) 466-6040  
 ; INFORMATION FOR SEQ ID NO: 40:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 23 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FRAGMENT TYPE: N-terminal  
 ; US-08-478-572-40

Query Match 89.5%; Score 51; DB 2; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.11;  
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 WTVRNSWDT 9  
 | | | | |  
 Db 14 WIVRNSWDT 22

RESULT 4  
 US-08-478-572-102  
 ; Sequence 102, Application US/08478572  
 ; Patent No. 5968526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Garman, Richard  
 ; APPLICANT: Greenstein, Julia  
 ; APPLICANT: Kuo, Mei-chang  
 ; APPLICANT: Rogers, Bruce  
 ; APPLICANT: Franzen, Henry  
 ; APPLICANT: Chen, Xian  
 ; APPLICANT: Evans, Sean  
 ; APPLICANT: Shaked, Ze'ev  
 ; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 ; TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 ; NUMBER OF SEQUENCES: 207

```
/
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
/ STREET: 610 LINCOLN STREET
/ CITY: WALTHAM
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02154
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/478,572
/ FILING DATE: 07-June-1995
/ CLASSIFICATION:
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/445,307
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CRAIG, ANNE I.
/ REGISTRATION NUMBER: 32,976
/ REFERENCE/DOCKET NUMBER: 017.6US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
/
/ INFORMATION FOR SEQ ID NO: 102:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 23 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: Peptide
/ FRAGMENT TYPE: N-terminal
/
/ US-08-478-572-102
/
/ Query Match 89.5%; Score 51; DB 2; Length 23;
/ Best Local Similarity 88.9%; Pred. No. 0.11;
/ Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/
/ Qy 1 WTVRNSWDT 9
/ Db 14 WIVRNSWDT 22
/
/ RESULT 5
/ US-08-484-296-40
/ Sequence 40, Application US/08484296
/ Patent No. 6268491
/ GENERAL INFORMATION:
/ APPLICANT: Garman, Richard
/ APPLICANT: Greenstein, Julia
/ APPLICANT: Kuo, Mei-chang
/ APPLICANT: Rogers, Bruce
/ APPLICANT: Franzen, Henry
/ APPLICANT: Chen, Xian
/ APPLICANT: Evans, Sean
/ APPLICANT: Shaked, Ze'ev
/ TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
/ FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
/
/ NUMBER OF SEQUENCES: 207
/ CORRESPONDENCE ADDRESSES:
/ ADDRESS: IMMULOGIC PHARMACEUTICAL CORPORATION
/ STREET: 610 LINCOLN STREET
/ CITY: WALTHAM
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02154
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/
/ CURRENT APPLICATION DATA:
```

```
/
/ APPLICATION NUMBER: US/08/484,296
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/445,307
/ FILING DATE: 07 June 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CRAIG, ANNE I.
/ REGISTRATION NUMBER: 32,976
/ REFERENCE/DOCKET NUMBER: 017.6US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
/
/ INFORMATION FOR SEQ ID NO: 40:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 23 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: Peptide
/ FRAGMENT TYPE: N-terminal
/
/ US-08-484-296-40
/
/ Query Match 89.5%; Score 51; DB 3; Length 23;
/ Best Local Similarity 88.9%; Pred. No. 0.11;
/ Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/
/ Qy 1 WTVRNSWDT 9
/ Db 14 WIVRNSWDT 22
/
/ RESULT 6
/ US-08-484-296-102
/ Sequence 102, Application US/08484296
/ Patent No. 6268491
/ GENERAL INFORMATION:
/ APPLICANT: Garman, Richard
/ APPLICANT: Greenstein, Julia
/ APPLICANT: Kuo, Mei-chang
/ APPLICANT: Rogers, Bruce
/ APPLICANT: Franzen, Henry
/ APPLICANT: Chen, Xian
/ APPLICANT: Evans, Sean
/ APPLICANT: Shaked, Ze'ev
/ TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
/ FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
/
/ NUMBER OF SEQUENCES: 207
/ CORRESPONDENCE ADDRESSES:
/ ADDRESS: IMMULOGIC PHARMACEUTICAL CORPORATION
/ STREET: 610 LINCOLN STREET
/ CITY: WALTHAM
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02154
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII TEXT
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/484,296
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/445,307
/ FILING DATE: 07 June 1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: CRAIG, ANNE I.
/ REGISTRATION NUMBER: 32,976
/ REFERENCE/DOCKET NUMBER: 017.6US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 466-6000
/ TELEFAX: (617) 466-6040
/
```

INFORMATION FOR SEQ ID NO: 102;  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 23 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: N-terminal  
 US-08-484-296-102

Query Match 89.5%; Score 51; DB 3; Length 23;  
 Best Local Similarity 88.9%; Pred. No. 0.11;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 DB 14 WIVRNSWDT 22

RESULT 7  
 US-08-482-142-25  
 ; Sequence 25, Application US/08482142  
 ; Patent No. 5820862

GENERAL INFORMATION:  
 APPLICANT: Garman, Richard  
 APPLICANT: Greenstein, Julia  
 APPLICANT: Kuo, Mei-chang  
 APPLICANT: Rogers, Bruce  
 APPLICANT: Franzen, Henry  
 APPLICANT: Chen, Xian  
 APPLICANT: Evans, Sean  
 APPLICANT: Shaked, Ze'ev  
 TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 NUMBER OF SEQUENCES: 207  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 STREET: 610 LINCOLN STREET  
 CITY: WALTHAM  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02154

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII TEXT  
 CURRENT APPLICATION DATA:  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/445,307  
 FILING DATE: 07 June 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CRAIG, ANNE I.  
 REGISTRATION NUMBER: 32,976  
 REFERENCE/DOCKET NUMBER: 017.6US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 466-6000  
 TELEFAX: (617) 466-6040  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 24 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: N-terminal  
 US-08-482-142-25

Query Match 89.5%; Score 51; DB 2; Length 24;  
 Best Local Similarity 88.9%; Pred. No. 0.12;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 DB 14 WIVRNSWDT 22

QY 1 WTVRNSWDT 9  
 DB 6 WIVRNSWDT 14

RESULT 8  
 US-08-482-142-88  
 ; Sequence 88, Application US/08482142  
 ; Patent No. 5820862

GENERAL INFORMATION:  
 APPLICANT: Garman, Richard  
 APPLICANT: Greenstein, Julia  
 APPLICANT: Kuo, Mei-chang  
 APPLICANT: Rogers, Bruce  
 APPLICANT: Franzen, Henry  
 APPLICANT: Chen, Xian  
 APPLICANT: Evans, Sean  
 APPLICANT: Shaked, Ze'ev  
 TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
 TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
 NUMBER OF SEQUENCES: 207  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
 STREET: 610 LINCOLN STREET  
 CITY: WALTHAM  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02154

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII TEXT  
 CURRENT APPLICATION DATA:  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/445,307  
 FILING DATE: 07 June 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CRAIG, ANNE I.  
 REGISTRATION NUMBER: 32,976  
 REFERENCE/DOCKET NUMBER: 017.6US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 466-6000  
 TELEFAX: (617) 466-6040  
 INFORMATION FOR SEQ ID NO: 88:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 24 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: N-terminal  
 US-08-482-142-88

Query Match 89.5%; Score 51; DB 2; Length 24;  
 Best Local Similarity 88.9%; Pred. No. 0.12;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRNSWDT 9  
 DB 6 WIVRNSWDT 14

RESULT 9  
 US-08-478-572-25  
 ; Sequence 25, Application US/08478572  
 ; Patent No. 5968526

GENERAL INFORMATION:  
 APPLICANT: Garman, Richard  
 APPLICANT: Greenstein, Julia  
 APPLICANT: Kuo, Mei-chang



APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017,6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-478-572-25  
Query Match 89.5%; Score 51; DB 2; Length 24;  
Best Local Similarity 88.9%; Pred. No. 0.12; Mismatches 0; Indels 1; Gaps 0;  
Matches 8; Conservative 0;  
QY 1 WTVRNSWDT 9  
DB 6 WIVRNSWDT 14  
RESULT 10  
US-08-478-572-88  
Sequence 88, Application US/08478572  
Patent No. 5968526  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA

COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,572  
FILING DATE: 07-June-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CRAIG, ANNE I.  
REGISTRATION NUMBER: 32,976  
REFERENCE/DOCKET NUMBER: 017,6US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 466-6000  
TELEFAX: (617) 466-6040  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 24 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-478-572-88  
Query Match 89.5%; Score 51; DB 2; Length 24;  
Best Local Similarity 88.9%; Pred. No. 0.12; Mismatches 0; Indels 1; Gaps 0;  
Matches 8; Conservative 0;  
QY 1 WTVRNSWDT 9  
DB 6 WIVRNSWDT 14  
RESULT 11  
US-08-484-296-25  
Sequence 25, Application US/08484296  
Patent No. 6268491  
GENERAL INFORMATION:  
APPLICANT: Garman, Richard  
APPLICANT: Greenstein, Julia  
APPLICANT: Kuo, Mei-chang  
APPLICANT: Rogers, Bruce  
APPLICANT: Franzen, Henry  
APPLICANT: Chen, Xian  
APPLICANT: Evans, Sean  
APPLICANT: Shaked, Ze'ev  
TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS  
TITLE OF INVENTION: FROM DERMATOPHAGOIDES (HOUSE DUST MITE)  
NUMBER OF SEQUENCES: 207  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
STREET: 610 LINCOLN STREET  
CITY: WALTHAM  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII TEXT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,296  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/445,307



```

;
; NUMBER OF SEQUENCES: 207
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
;
; US-08-482-142-39

```

```

Query Match      89.5%; Score 51; DB 2; Length 29;
Best Local Similarity 88.9%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 WTVRNSWDT 9
Db      14 WIVRNSWDT 22

```

```

RESULT 15
US-08-482-142-101
; Sequence 101, Application US/08482142
; Patent No. 5820862
;
; GENERAL INFORMATION:
; APPLICANT: Garman, Richard
; APPLICANT: Greenstein, Julia
; APPLICANT: Kuo, Mei-chang
; APPLICANT: Rogers, Bruce
; APPLICANT: Franzen, Henry
; APPLICANT: Chen, Xian
; APPLICANT: Evans, Sean
; APPLICANT: Shaked, Ze'ev
; TITLE OF INVENTION: T CELL EPITOPES OF THE MAJOR ALLERGENS
; NUMBER OF SEQUENCES: 207
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,142
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,307
; FILING DATE: 07 June 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG, ANNE I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 017.6US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
;
; US-08-482-142-101

```

```

Query Match      89.5%; Score 51; DB 2; Length 29;
Best Local Similarity 88.9%; Pred. No. 0.14;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 WTVRNSWDT 9
Db      14 WIVRNSWDT 22

```

```

Search completed: October 27, 2004, 17:53:01
Job time : 3.19124 secs

```

10

11

12

13

14

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:40:03 ; Search time 39 Seconds  
(without alignments)  
22,204 Million cell updates/sec

Title: US-09-867-159A-5

Perfect score: 57

Sequence: 1 WTVRNSWD 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 791

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match	Length	DB ID	Description
1	26	45.6	9	2	PT0299
2	22	38.6	5	2	PT0580
3	22	38.6	8	2	A38887
4	19	33.3	9	2	A24244
5	18	31.6	6	2	A31263
6	17	29.8	6	2	PT0519
7	17	29.8	7	2	S33244
8	17	29.8	7	2	S33245
9	17	29.8	7	2	S33246
10	16	28.1	4	2	PT0661
11	16	28.1	6	2	A61068
12	16	28.1	7	2	PH1602
13	16	28.1	7	4	I53382
14	16	28.1	8	2	S3288
15	16	28.1	8	2	JS0315
16	16	28.1	8	2	JS0316
17	16	28.1	8	2	JS0317
18	16	28.1	9	2	A43848
19	15	26.3	6	2	B31263
20	15	26.3	7	2	PN0649
21	15	26.3	7	2	A61081
22	15	26.3	8	2	S10596
23	15	26.3	8	2	JS0318
24	15	26.3	9	2	S07205
25	15	26.3	9	2	S07204
26	15	26.3	9	2	PT0231
27	15	26.3	9	2	I58350
28	15	26.3	9	2	PC2021
29	15	26.3	9	2	D57444

30 14 24.6 5 2 PT0308 Ig heavy chain CRD  
31 14 24.6 6 2 S66195 alcohol dehydrogen  
32 14 24.6 9 2 A28495 conopressin G - co  
33 14 24.6 9 2 B28495 conopressin S - co  
34 14 24.6 9 2 S07241 litorin - Rohde's  
35 14 24.6 9 2 PT0270 Ig heavy chain CRD  
36 14 24.6 9 2 S39040 lysine-conopressin  
37 13 22.8 5 2 PT0590 T-cell receptor be  
38 13 22.8 6 2 B35840 cerebellar degener  
39 13 22.8 6 2 PT0830 T-cell receptor be  
40 13 22.8 6 2 PD0028 rev-kinin 2 - pena  
41 13 22.8 7 2 S09066 globulin IV alpha  
42 13 22.8 8 2 B4960 neuropeptide Ied-C  
43 13 22.8 8 2 S08996 hypertrehalosemic  
44 13 22.8 8 2 B49823 adipokinetic hormo  
45 13 22.8 8 2 A59495 vesicle associated

#### ALIGNMENTS

##### RESULT 1

PT0299

Ig heavy chain CRD3 region (clone 5-103B) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C:Accession: PT0299

R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j

A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0299

A:Molecule type: DNA

A:Residues: 1-9 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 45.6%; Score 26; DB 2; Length 9;

Best Local Similarity 80.0%; Pred.No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RNSWD 8

Db 1 RESWD 5

##### RESULT 2

PT0580

T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C:Accession: PT0580

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A:Reference number: PT0509; MUID:91277601; PMID:1711558

A:Accession: PT0580

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FEE>

A:Experimental source: day 19 fetal thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match 38.6%; Score 22; DB 2; Length 5;

Best Local Similarity 75.0%; Pred.No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NSWD 8

Db 2 SSWD 5

##### RESULT 3

A38887  
T-cell receptor gamma chain (St.2) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999  
C:Accession: A38887  
R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.  
Mol. Cell. Biol. 11, 5902-5909, 1991  
A>Title: Rearrangement and functional-site sequence analyses of T-cell receptor gamma gene  
A:Reference number: A41946; MUID:92049316; PMID:1658619  
A:Accession: A38887  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-8 <WHE>  
C:Keywords: T-cell receptor

Query Match 38.6%; Score 22; DB 2; Length 8;  
Best Local Similarity 75.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SWDT 9  
| | |  
Db 2 SWDS 5

RESULT 4  
adipokinetin hormone - bollworm  
N:Alternate names: Hez-AKH  
C:Species: Heliothis zea (bollworm, corn earworm, tomato fruitworm)  
C>Date: 31-Mar-1998 #sequence\_revision 23-Mar-1995 #text\_change 09-Jul-2004  
C:Accession: A24244  
R:Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Holman, G.M.; Wagner, R.M.; Ridgway  
Biochem. Biophys. Res. Commun. 135, 622-628, 1996  
A>Title: Isolation and primary structure of a peptide from the corpora cardiaca of Helic  
A:Reference number: A24244; MUID:86186794; PMID:3964263  
A:Accession: A24244  
A:Molecule type: protein  
A:Residues: 1-9 <UAP>  
A:Cross-references: UNIPROT:P08901  
C:Superfamily: adipokinetin hormone  
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:9/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 33.3%; Score 19; DB 2; Length 9;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TVRNSW 7  
| | |  
Db 3 TFTSSW 8

RESULT 5  
A31263  
dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - Plasmodium f  
C:Species: Plasmodium falciparum  
C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
C:Accession: A31263  
R:Peterson, D.S.; Walliker, D.; Wellens, T.E.  
Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
A>Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase  
A:Reference number: A94217; MUID:89057886; PMID:2904149  
A:Accession: A31263  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-6 <PET>  
C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match 31.6%; Score 18; DB 2; Length 6;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 6 SWDT 9  
| | |  
Db 2 SWES 5

RESULT 6  
PT0519  
T-cell receptor beta chain V-D-J region (100-4C) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C:Accession: PT0519  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A>Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0519  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-6 <FEE>  
A:Experimental source: adult thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 29.8%; Score 17; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WD 8  
| |  
Db 5 WD 6

RESULT 7  
S33244  
neuromodulatory peptide Wamide-1 - giant African snail  
C:Species: Achatina fulica (giant African snail)  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S33244  
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
FEBS Lett. 323, 104-108, 1993  
A>Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of th  
A:Reference number: S33244; MUID:93265912; PMID:8495720  
A:Accession: S33244  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIN>  
A:Cross-references: UNIPROT:P35921

Query Match 29.8%; Score 17; DB 2; Length 7;  
Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WTVNSW 7  
| |  
Db 1 WKMSVW 7

RESULT 8  
S33245  
neuromodulatory peptide Wamide-2 - giant African snail  
C:Species: Achatina fulica (giant African snail)  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S33245  
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
FEBS Lett. 323, 104-108, 1993  
A>Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of th  
A:Reference number: S33244; MUID:93265912; PMID:8495720  
A:Accession: S33245  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIN>  
A:Cross-references: UNIPROT:P35919

Query Match 29.8%; Score 17; DB 2; Length 7;

Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
| : |  
Db 1 WREMSVW 7

RESULT 9  
S33246  
neuromodulatory peptide Wamide-3 - giant African snail  
C:Species: Achatina fulica (giant African snail)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S33246  
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.  
PDBS Lett. 323; 104-108, 1993  
A:Title: Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of *Achatina fulica*  
A:Reference number: S33244; MUID:93265912; PMID:9495720  
A:Accession: S33246  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-7 <MIN>  
A:Cross-references: UNIPROT:P35920

Query Match 29.8%; Score 17; DB 2; Length 7;  
Best Local Similarity 28.6%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
| : |  
Db 1 WKQMSVW 7

RESULT 10  
PT0661  
T-cell receptor beta chain V-D-J region (121-1BV) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 09-Jul-2004  
C:Accession: PT0661  
R:Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A:Reference number: PT0509; MUID:91277601; PMID:1711558  
A:Accession: PT0661  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-4 <FE>  
A:Cross-references: UNIPROT:Q8BZ07; UNIPROT:Q8CCN5  
A:Experimental source: day 4 postnatal thymus, strain BALB/c  
C:Keywords: T-cell receptor

Query Match 28.1%; Score 16; DB 2; Length 4;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WT 2  
| |  
Db 3 WT 4

RESULT 11  
A61068  
locustakinin - migratory locust  
C:Species: Locusta migratoria (migratory locust)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: A61068  
R:Schoofs, L.; Holman, G.M.; Proost, P.; Van Damme, J.; Hayes, T.K.; De Loof, A.  
Regul. Pept. 37, 49-57, 1992  
A:Title: Locustakinin, a novel myotrophic peptide from Locusta migratoria, isolation, purification and synthesis  
A:Reference number: A61068; MUID:92262851; PMID:1585017  
A:Accession: A61068  
A:Molecule type: protein  
A:Residues: 1-6 <SCH>

A:Cross-references: UNIPROT:P41491  
C:Keywords: amidated carboxyl end; cephalomyotropic peptide; neuropeptide  
F:6/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.1%; Score 16; DB 2; Length 6;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 NSW 7  
| | |  
Db 3 SSW 5

RESULT 12  
PH1602  
Ig H chain V-D-J region (wild-type clone 313) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jun-1994 #sequence\_revision 02-Jun-1994 #text\_change 17-Mar-1999  
C:Accession: PH1602  
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.  
J. Exp. Med. 178, 317-329, 1993  
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice  
A:Reference number: PH1580; MUID:93301609; PMID:8315387  
A:Accession: PH1602  
A:Molecule type: DNA  
A:Residues: 1-7 <LEV>  
A:Experimental source: bone marrow pre-B lymphocyte  
C:Keywords: immunoglobulin

Query Match 28.1%; Score 16; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WT 2  
| |  
Db 6 WT 7

RESULT 13  
I55382  
hypothetical peptide PA11 promoter region - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 16-Apr-1999 #sequence\_revision 16-Apr-1999 #text\_change 20-Apr-2000  
C:Accession: I55382  
R:Dawson, S.J.; Wiman, B.; Hamsten, A.; Green, F.; Humphries, S.; Henney, A.M.  
J. Biol. Chem. 268, 10739-10745, 1993  
A:Title: The two allele sequences of a common polymorphism in the promoter of the plasminogen activator gene  
A:Reference number: I55382; MUID:93266509; PMID:8388372  
A:Accession: I55382  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-7 <DAW>  
A:Cross-references: GB:M91557; NID:gi90020; PIDN:AAA60110.1; PID:gi90021  
C:Comment: This is the hypothetical translation of a sequence from the PA11 gene promoter:  
C:Genetics:  
A:Gene: GDB:PA11  
A:Cross-references: GDB:120297; OMIM:173360  
A:Map position: 7q21.3-7q22

Query Match 28.1%; Score 16; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WT 2  
| |  
Db 1 WT 2

RESULT 14  
SI9288  
acylase - Kluyvera cryocrescens  
C:Species: Kluyvera cryocrescens  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S19288  
R;Martin, J.; Slade, A.; Aitken, A.; Arche, R.; Virden, R.  
Biochem. J. 280, 659-662, 1991  
A;Title: Chemical modification of serine at the active site of penicillin acylase from K  
A;Reference number: S19288; MUID: 92109664; PMID: 1764029  
A;Accession: S19288  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-8 <MAR>  
A;Cross-references: UNIPROT:Q7M124

Query Match 28.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 NSW 7  
|  
|  
Db 2 NMW 4

## RESULT 15

JS0315  
Leucokinin V - Madeira cockroach  
C;Species: Leucophaea maderae (Madeira cockroach)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
C;Accession: JS0315  
R;Holman, G.M.; Cook, B.J.; Nachman, R.J.  
Comp. Biochem. Physiol. C 88, 27-30, 1987  
A;Title: Isolation, primary structure, and synthesis of leucokinins V and VI: myotropic  
A;Reference number: JS0315  
A;Accession: JS0315  
A;Molecule type: protein  
A;Residues: 1-8 <HOL>  
A;Cross-references: UNIPROT:P19987  
C;Comment: Leucokinins, a family of cephalomyotropic peptides, stimulate contractile act  
C;Keywords: amidated carboxyl end; cephalomyotropic peptide  
F;8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 28.1%; Score 16; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 NSW 7  
|  
|  
Db 5 SSW 7

Search completed: October 27, 2004, 18:55:05  
Job time : 42 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:39:47 ; Search time 187 Seconds  
(without alignments)  
27.692 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1598

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27.5	48.2	9	Q38366	Q38366 bacteriophage
2	25	43.3	9	Q9H3Y3	Q9H3Y3 homo sapien
3	21	36.8	4	OCF3_OCTMI	P58649 octopus min
4	21	36.8	7	Q8GL04	Q8GL04 borrelia bu
5	21	36.8	8	LCK1_LEUMA	P21140 leucophaea
6	21	36.8	8	LCK3_LEUMA	P21142 leucophaea
7	21	36.8	8	PK1_PERAM	P82685 periplaneta
8	21	36.8	8	PK3_PERAM	P82687 periplaneta
9	21	36.8	8	PK5_PERAM	P82689 periplaneta
10	21	36.8	8	Q8G540	Q8G540 borrelia bu
11	21	36.8	8	Q8GL26	Q8GL26 borrelia bu
12	21	36.8	9	Q9H3Y3	Q9H3Y3
13	19	33.3	8	Q37854	Q37854 bacteriophage
14	19	33.3	8	Q8GL21	Q8GL21 borrelia bu
15	19	33.3	9	Q8GL31	Q8GL31 borrelia bu
16	18	31.6	8	C125_CVPDO	P83661 cyphononyx
17	18	31.6	8	AAB33374	AAB33374 homo sapi
18	18	31.6	9	AAB33374	AAB33374 human imm
19	17	29.8	7	WNA1_ACHFU	P35919 achatina fu
20	17	29.8	7	WNA3_ACHFU	P35920 achatina fu
21	17	29.8	7	WNA3_ACHFU	P35921 achatina fu
22	17	29.8	7	Q49223	Q49223 glycine max
23	17	29.8	8	Q62721	Q62721 rattus norv
24	17	29.8	8	P79940	P79940 xenopus lae
25	17	29.8	9	Q6LDB5	Q6LDB5 homo sapien
26	17	29.8	9	Q801K0	Q801K0 illicura mil
27	17	29.8	9	Q801K1	Q801K1 chiroiphia
28	17	29.8	9	Q801K2	Q801K2 chiroiphia
29	17	29.8	9	AAB27101	AAB27101 homo sapi
30	16	28.1	6	LCK1_LOCOMI	P41491 locusta mig
31	16	28.1	8	LCK2_LEUMA	P21141 leucophaea

32 16 28.1 8 1 LCK4\_LEUMA P21143 leucophaea  
33 16 28.1 8 1 LCK5\_LEUMA P19987 leucophaea  
34 16 28.1 8 1 LCK6\_LEUMA P19988 leucophaea  
35 16 28.1 8 1 LCK7\_LEUMA P19989 leucophaea  
36 16 28.1 8 1 PK2\_PERAM P82686 periplaneta  
37 16 28.1 8 1 PK4\_PERAM P82688 periplaneta  
38 16 28.1 8 1 PIP\_BRANA P81707 brassica na  
39 16 28.1 8 2 Q7ML24 Q7ML24 kluyvera ci  
40 16 28.1 8 2 Q99MN0 Q99MN0 mus musculu  
41 16 28.1 9 2 Q7LDX2 Q7LDX2 urostrophus  
42 16 28.1 9 2 Q6VCX0 Q6VCX0 streptomyce  
43 16 28.1 9 2 Q9RSM1 Q9RSM1 staphylococ  
44 16 28.1 9 2 Q90350 Q90350 gb virus c/  
45 16 28.1 9 2 AAQ22766 AAQ22766 streptomy

#### ALIGNMENTS

RESULT 1  
Q38366 PRELIMINARY; PRT; 9 AA.  
AC Q38366;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE E Gene product (Fragment).  
OS Bacteriophage phi-X174.  
OC Viruses; ssDNA viruses; Microviridae; Microvirus.  
OX NCBI\_TaxID=10847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINS=88118956; PubMed=2963134;  
RA Buckley K.J., Hayashi M.;  
RT Role of premature translational termination in the regulation of  
RL expression of the phi X174 lysis gene."  
RL J. Mol. Biol. 198:599-607(1987).  
DR EMBL; X07809; CAA30668.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1207 MW; C093B37731B36412 CRC64;

Query Match 48.2%; Score 27.5; DB 2; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.8e+06;  
Matches 5; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 1 WTVRNSWDT 9  
|||  
Db 4 WTL---WDT 9

RESULT 2  
Q9H3Y3 PRELIMINARY; PRT; 9 AA.  
AC Q9H3Y3;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE DJ839B11.1 (Novel protein with a Kunitz/Bovine pancreatic trypsin  
inhibitor domain and WAP-type (Whey Acidic Protein) 'four-disulfide  
core' domains) (Fragment).  
DE Core 'dj461p17.1';  
GN Name=dj461p17.1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lloyd D.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL121778; CAB76844.1; -.  
FT NON TER 9  
SQ SEQUENCE 9 AA; 1036 MW; 2C417B01B412D1B3 CRC64;

Query Match 43.9%; Score 25; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVR 4  
 ||||  
 DB 2 WTVR 5

RESULT 3  
 OCP3\_OCTMI STANDARD; PRT; 4 AA.  
 AC P58649;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Cardioactive peptides Ocp-3/Ocp-4.  
 OS Octopus minor (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE=Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.;  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 RT Octopus minor.";  
 RL Peptides 21:623-630(2000).  
 CC -! FUNCTION: Cardioactive; has both positive chronotropic and  
 CC inotropic effects on the heart. Ocp-4 is a 1000 time less active  
 CC than Ocp-3.  
 CC -! SUBCELLULAR LOCATION: Secreted.  
 CC -! PTM: Ocp-4 has D-Ser instead of L-Ser.  
 CC -! MASS SPECTROMETRY: MW=395.2; METHOD=MALDI; RANGE=1-4; NOTE=Ref.1.  
 KW D-amino acid; Direct protein sequencing; Hormone.  
 FT MOD\_RES 2  
 FT SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;  
 SQ

Query Match 36.8%; Score 21; DB 1; Length 4;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SWD 8  
 ||||  
 DB 2 SWD 4

RESULT 4  
 O8GL04 PRELIMINARY; PRT; 7 AA.  
 AC O8GL04;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE PF-50 protein (Fragment).  
 GN Name=PF-50;  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Plasmid group cp32-5.  
 OC Bacteriia; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=93-0107;  
 RX MEDLINE=22610300; PubMed=12724373;  
 RA Miller J.C., Stevenson B.;  
 RT "Immunological and genetic characterization of Borrelia burgdorferi  
 RT BapA and BppA proteins.";  
 RL Microbiology 149:1113-1125(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=93-0107;  
 RX MEDLINE=22590544; PubMed=14629041;

RA Stevenson B., Miller J.C.;  
 RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
 RT erp genes generates sequence identity amidst diversity.";  
 RL J. Mol. Evol. 57:309-324(2003).  
 DR EMBL; AY142103; AANI7848.1; -;  
 KW Plasmid.  
 FT NON\_TER. 1  
 SQ SEQUENCE 7 AA; 914 MW; 6337244330504310 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 7;  
 Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVR 5  
 ||||  
 DB 1 WIKN 5

RESULT 5  
 LCK1\_LEUMA STANDARD; PRT; 8 AA.  
 ID LCK1\_LEUMA  
 AC P21140;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Leucokinin I (L-I).  
 OS Leucophaea maderae (Madeira cockroach).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 CC Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of two neuropeptides from  
 RT Leucophaea maderae: members of a new family of Cephalomyotopins.";  
 RL Comp. Biochem. Physiol. 84C:205-211(1986).  
 CC -! FUNCTION: This cephalomyotropic peptide stimulates contractile  
 CC activity of cockroach proctodeum (hindgut).  
 CC -! SUBCELLULAR LOCATION: Secreted.  
 CC Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 8  
 FT SEQUENCE 8 AA; 893 MW; DC6365B449CDC76A CRC64;  
 SQ

Query Match 36.8%; Score 21; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 NSW 7  
 ||||  
 DB 5 NSW 7

RESULT 6  
 LCK3\_LEUMA STANDARD; PRT; 8 AA.  
 ID LCK3\_LEUMA  
 AC P21142;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Leucokinin III (L-III).  
 OS Leucophaea maderae (Madeira cockroach).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 CC Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Primary structure and synthesis of two additional neuropeptides from  
 RT Leucophaea maderae: members of a new family of Cephalomyotopins.";

RL Comp. Biochem. Physiol. 84C:271-276(1986).  
CC -!- FUNCTION: This cephalomyotropic peptide stimulates contractile  
CC activity of cockroach prothodum (hindgut).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 910 MW; DC6365B449C866DA CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 NSW 7  
Db |||  
5 NSW 7  
  
RESULT 7  
PK3\_PERAM STANDARD; PRT; 8 AA.  
AC P82685;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinin-1 (Pea-K-1).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;  
OC Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.  
RC TISSUE=Corpora cardiaca;  
RX MEDLINE=98010462; PubMed=9350979;  
RA Fredel R., Kellner R., Rapus J., Penzlin H., Gade G.;  
RT "Isolation and structural elucidation of eight kinins from the  
RT retrocerebral complex of the American cockroach, Periplaneta  
RT americana.";  
RL Regul. Pept. 71:199-205(1997).  
CC -!- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
CC activity).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MASS SPECTROMETRY: MW=949.26; METHOD=Electrospray; RANGE=1-8;  
CC NOTE=Ref.1.  
CC -!- SIMILARITY: Belongs to the kinin family.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 950 MW; 326365B449D5A774 CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 NSW 7  
Db |||  
5 NSW 7  
  
RESULT 8  
PK3\_PERAM STANDARD; PRT; 8 AA.  
AC P82687;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinin-3 (Pea-K-3).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;  
OC Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.

RC TISSUE=Corpora cardiaca;  
RX MEDLINE=98010462; PubMed=9350979;  
RA Fredel R., Kellner R., Rapus J., Penzlin H., Gade G.;  
RT "Isolation and structural elucidation of eight kinins from the  
RT retrocerebral complex of the American cockroach, Periplaneta  
RT americana.";  
RL Regul. Pept. 71:199-205(1997).  
CC -!- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
CC activity).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MASS SPECTROMETRY: MW=907.92; METHOD=Electrospray; RANGE=1-8;  
CC NOTE=Ref.1.  
CC -!- SIMILARITY: Belongs to the kinin family.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 909 MW; DC6365B449D5A76A CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 NSW 7  
Db |||  
5 NSW 7  
  
RESULT 9  
PK5\_PERAM STANDARD; PRT; 8 AA.  
AC P82689;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Kinin-5 (Pea-K-5).  
OS Periplaneta americana (American cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;  
OC Blattidae; Periplaneta.  
OX NCBI\_TaxID=6978;  
RN [1]  
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.  
RC TISSUE=Corpora cardiaca;  
RX MEDLINE=98010462; PubMed=9350979;  
RA Fredel R., Kellner R., Rapus J., Penzlin H., Gade G.;  
RT "Isolation and structural elucidation of eight kinins from the  
RT retrocerebral complex of the American cockroach, Periplaneta  
RT americana.";  
RL Regul. Pept. 71:199-205(1997).  
CC -!- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
CC activity).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- MASS SPECTROMETRY: MW=864.10; METHOD=Electrospray; RANGE=1-8;  
CC NOTE=Ref.1.  
CC -!- SIMILARITY: Belongs to the kinin family.  
KW Amidation; Direct protein sequencing; Neuropeptide.  
FT MOD\_RES 8 8 Glycine amide.  
SQ SEQUENCE 8 AA; 865 MW; C76365B449CDC775 CRC64;  
  
Query Match 36.8%; Score 21; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 NSW 7  
Db |||  
5 NSW 7  
  
RESULT 10  
Q8G940 PRELIMINARY; PRT; 8 AA.  
ID Q8G940  
AC Q8G940;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE PF-50 protein (Fragment).

GN Name=PF-50;  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OG Plasmid group cp32-9, and plasmid group cp32-12.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
[1]

RN NAME=PF-50;  
RP SEQUENCE FROM N.A.  
RC STRAIN=DN127c19-2; PLASMID=group cp32-9;  
RX MEDLINE=22610300; PubMed=12724373;  
RA Miller J.C., Stevenson B.;

RT "Immunological and genetic characterization of Borrelia burgdorferi  
BapA and EppA proteins.";  
RL Microbiology 149:1113-1125 (2003).  
[2]

RP SEQUENCE FROM N.A.  
RC STRAIN=DN127c19-2, and Sh-2-82;  
RX PLASMID=group cp32-9, and group cp32-12;  
RX MEDLINE=22990544; PubMed=14629041;

RA Stevenson B., Miller J.C.;  
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
erp genes generates sequence identity amidst diversity.";

RL J. Mol. Evol. 57:309-324 (2003).  
DR EMBL; AY142104; AAN17853.1; -;  
DR EMBL; AY142097; AAN17907.1; -;  
KW Plasmid.  
FT NON\_TER

SQ SEQUENCE 8 AA; 1042 MW; 1437244330504373 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 8;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5  
| : : |  
Db 2 WIKN 6

RESULT 11

O8GL26 PRELIMINARY; PRT; 9 AA.

ID O8GL26;  
AC O8GL26;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE PF-50 protein (Fragment).  
GN Name=PF-50;  
OS Borrelia burgdorferi (Lyme disease spirochete).

OG Plasmid group cp32-5.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
[1]

RN NAME=PF-50;  
RP SEQUENCE FROM N.A.

RC STRAIN=Sh-2-82;  
RX MEDLINE=22990544; PubMed=14629041;

RA Stevenson B., Miller J.C.;  
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
erp genes generates sequence identity amidst diversity.";

RL J. Mol. Evol. 57:309-324 (2003).  
DR EMBL; AY142092; AAN17873.1; -;  
KW Plasmid.  
FT NON\_TER

SQ SEQUENCE 9 AA; 1206 MW; 5A4A244330504373 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 9;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5  
| : : |  
Db 3 WIKN 7

RESULT 12

O9R9C4 PRELIMINARY; PRT; 9 AA.

ID O9R9C4;  
AC O9R9C4;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Borrelia burgdorferi plasmid cp32-2, possible partition proteins,  
complete cds (PF-50 protein) (Fragment).  
GN Name=PF-50;  
OS Borrelia burgdorferi (Lyme disease spirochete).

OG Plasmid cp32-2, and plasmid group cp32-4.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
[1]

RN NAME=PF-50;  
RP SEQUENCE FROM N.A.

RC STRAIN=B31; PLASMID=cp32-2;  
RX MEDLINE=98361033; PubMed=9695920;

RA Stevenson B., Casjens S., Rosa P.;  
RT "Evidence of past recombination events among the genes encoding the  
erp antigens of Borrelia burgdorferi.";

RL Microbiology 144:1869-1879 (1998).  
[2]

RP SEQUENCE FROM N.A.  
RC STRAIN=Sh-2-82; PLASMID=group cp32-4;  
RX MEDLINE=22990544; PubMed=14629041;

RA Stevenson B., Miller J.C.;  
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
erp genes generates sequence identity amidst diversity.";

RL J. Mol. Evol. 57:309-324 (2003).  
DR EMBL; AF022479; AAC35438.1; -;  
DR EMBL; AY142091; AAN17865.1; -;  
KW Plasmid.  
FT NON\_TER

SQ SEQUENCE 9 AA; 1155 MW; 4E1A244330504373 CRC64;

Query Match 36.8%; Score 21; DB 2; Length 9;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5  
| : : |  
Db 3 WIKN 7

RESULT 13

O37854 PRELIMINARY; PRT; 8 AA.

ID O37854;  
AC O37854;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Coliphage gene of unknown function, 5'end. (Fragment).  
OS Bacteriophage R17.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;  
OC Levivirus.

RN NAME=PF-50;  
RP SEQUENCE FROM N.A.

RC STRAIN=Sh-2-82;  
RX MEDLINE=73224987; PubMed=4352721;

RA Rensing U.P.E.;  
RT "A sequence of seventy-three nucleotides from the coliphage R17  
genome.";

RL Biochem. J. 131:593-604 (1973).  
DR EMBL; M24820; AAA72755.1; -;  
KW Plasmid.  
FT NON\_TER

SQ SEQUENCE 8 AA; 969 MW; ECB45412C1E72726 CRC64;

Query Match 33.3%; Score 19; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: October 27, 2004, 18:54:20  
Job time : 189 secs

QY 3 VRNS 6  
DB 5 VRNS 8

# RESULT 14

Q8GL21 Q8GL21 PRELIMINARY; PRT; 8 AA.  
AC Q8GL21;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE PF-50 protein (Fragment).  
GN Name=PF-50;  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OG Plasmid group cp32-8.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sh-2-82;  
RX MEDLINE=22990544; PubMed=14629041;  
RA Stevenson B., Miller J.C.;  
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
RT erp genes generates sequence identity amidst diversity.";  
EL J. Mol. Evol. 57:309-324(2003).  
DR EMBL; AY142094; AAN17903.1; -.  
KW Plasmid.  
FT NON TER 1  
SQ SEQUENCE 8 AA; 1042 MW; 1437244337204373 CRC64;  
Query Match 33.3%; Score 19; DB 2; Length 8;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5  
DB 2 WILKN 6

# RESULT 15

Q8GL31 Q8GL31 PRELIMINARY; PRT; 9 AA.  
AC Q8GL31;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE PF-50 protein (Fragment).  
GN Name=PF-50;  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OG Plasmid group cp32-1.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sh-2-82;  
RX MEDLINE=22990544; PubMed=14629041;  
RA Stevenson B., Miller J.C.;  
RT "Intra- and interbacterial genetic exchange of Lyme disease spirochete  
RT erp genes generates sequence identity amidst diversity.";  
EL J. Mol. Evol. 57:309-324(2003).  
DR EMBL; AY142089; AAN17869.1; -.  
KW Plasmid.  
FT NON TER 1  
SQ SEQUENCE 9 AA; 1206 MW; 5A4A244337204373 CRC64;

Query Match 33.3%; Score 19; DB 2; Length 9;  
Best Local Similarity 40.0%; Pred. No. 1.8e+06;  
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVRN 5  
DB 3 WILKN 7



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:35:41 ; Search time 150 Seconds  
(without alignments)  
21,524 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWD 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 313949

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	57	100.0	9	5	AAO20571	Cysteine
2	57	100.0	9	5	AB98536	Cysteine
3	41	71.9	8	5	ABJ04085	C.parvum
4	37	64.9	8	2	AAR77352	Cysteine
5	36	63.2	9	4	AAB76056	Trypanoso
6	33	57.9	9	2	AAW60399	Tumour ho
7	33	57.9	9	2	AAW93726	Mouse B16
8	33	57.9	9	3	AAB21816	Murine me
9	33	57.9	9	4	AAE06394	Tumour ho
10	30	52.6	5	4	AAB85112	C. parvum
11	30	52.6	5	5	ABJ04086	C.parvum
12	30	52.6	5	8	ADQ95813	Cryptospor
13	29	50.9	7	8	ADH12805	Abalone c
14	29	50.9	8	4	AAB85131	P. vincke
15	29	50.9	8	8	ADQ95812	Cryptospor
16	29	50.9	9	8	ADK10416	Human pap
17	23	50.9	9	8	ADK10223	Human pap
18	28	49.1	9	7	ADM18218	Chlamydia
19	27	47.4	6	2	AAR57079	Fasciola
20	27	47.4	9	5	AAU71234	Human MHC
21	26	45.6	7	5	ABG60349	Selective
22	26	45.6	8	7	ADL17323	Human scr
23	26	45.6	9	4	ADL17323	Human scr
24	26	45.6	9	5	ABG56538	Phage clo
25	26	45.6	9	5	ABG97285	Human.leu
					ABJ16457	Zinc tran

26	26	45.6	9	5	ABJ15476	Zinc tran
27	26	45.6	9	6	ABJ38076	Human cyt
28	26	45.6	9	8	ADN64445	HLA bindi
29	26	45.6	9	8	ADP30382	Human sec
30	25	43.9	7	2	AAW30429	HRE-I aff
31	25	43.9	8	3	AAJ77698	TSG-6 bin
32	25	43.9	9	4	AAU24450	Human MHC
33	25	43.9	9	4	AAU24036	Human MHC
34	25	43.9	9	4	AAU23831	Human MHC
35	25	43.9	9	4	AAU24154	Human MHC
36	25	43.9	9	5	ABG34271	Human leu
37	25	43.9	9	7	ADK23765	Human 98p
38	25	43.9	9	8	ADK02690	Hepatitis
39	25	43.9	9	8	ADK07503	Hepatitis
40	24	42.1	7	2	AAJ30217	Physarum
41	24	42.1	7	7	ADC54100	Commam am
42	24	42.1	7	7	ADL17066	Phage-dis
43	24	42.1	8	4	AAB78541	HIV-2 gp
44	24	42.1	8	5	ADE03135	Hybrid po
45	24	42.1	8	7	ADG14433	Transport

ALIGNMENTS

RESULT 1  
AAO20571  
ID AAO20571 standard; peptide; 9 AA.

AC AAO20571;  
XX  
DT 02-JAN-2003 (first entry)  
XX  
DE Cysteine protease epitope peptide region, SEQ ID No 5.

XX Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; cysteine protease protein; enzyme;  
KW atopic eczema; epitope.

XX OS Dermatophagoides pteronyssinus.

XX PN WO200278736-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-FR001098.

XX PR 30-MAR-2001; 2001FR-00004370.

XX PR 03-MAY-2001; 2001FR-00005929.

XX PR 29-MAY-2001; 2001US-00867159.

XX PA (ANTI-) ANTIALIS SARL.

XX PI Loxia E, Terrasse G, Trehin Y;

XX DR WPI, 2002-750636/81.

XX PT Antiallergic compositions containing an anti-histamine, a histamine

XX PT synthesis inhibitor, and optionally an allergen or nucleic acid coding

XX PT for the allergen.

XX PS Claim 14; Page 11; 32pp; French.

XX CC The invention relates to antiallergic compositions containing an anti-histamine, a histamine synthesis inhibitor, and optionally an allergen or nucleic acid coding sequence coding for the allergen, together with a pharmaceutical carrier. The pharmaceutical composition of the invention is useful as a non-specific antiallergic treatment, and also useful in the treatment of allergic hypersensitivity, allergic asthma, allergic rhinitis, and allergic atopic eczema. This sequence represents a peptide of a cysteine protease epitope region relating to the antiallergic

CC compositions of the invention

XX Sequence 9 AA;

Query Match 100.0%; Score 57; DB 5; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9

DB 1 WTVNSWDT 9

RESULT 2

ABB98536

ID ABB98536 standard; peptide; 9 AA.

XX

AC ABB98536;

DT 13-DEC-2002 (first entry)

XX

DE Cysteine protease epitope #3.

XX

KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;

KW immunotherapy; allergen; allergic hypersensitivity reaction;

KW allergic asthma; allergic rhinitis; allergic atopic eczema;

KW cysteine protease.

XX

OS Dermatophagoides pteronyssinus.

XX

PN FR2822709-A1.

XX

PD 04-OCT-2002.

XX

PF 03-MAY-2001; 2001FR-00005929.

XX

PR 30-MAR-2001; 2001FR-00004370.

XX

PA (ANTI-) ANTIALIS SARL.

XX

PI Loria E, Terrasse G, Trehin Y;

XX

DR WPI; 2002-735037/80.

XX

PT Antiallergic composition, useful for preventing and treating e.g. asthma,

PT rhinitis or eczema, containing at least two of allergen, antihistamine

PT and histamine synthesis inhibitor.

XX

PS Claim 8; Page 6; 33pp; French.

XX

CC The present invention relates to an antiallergic pharmaceutical  
CC composition (I) comprising a pharmaceutical carrier containing an active  
CC agent combination of at least two of: an allergen; an antihistamine; and  
CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
CC rhinitis or allergic atopic eczema, in babies, children or adults. The  
CC present sequence is a peptide fragment (epitope) of cysteine protease  
CC from Dermatophagoides pteronyssinus, which was used as an allergen in the  
CC invention

XX Sequence 9 AA;

Query Match

Best Local Similarity 100.0%; Score 57; DB 5; Length 9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVNSWDT 9

DB 1 WTVNSWDT 9

RESULT 3

ABJ04065

ID ABJ04065 standard; peptide; 8 AA.

XX

AC ABJ04065;

DT 27-SEP-2002 (first entry)

XX

DE C parvum cryptopain peptide fragment SEQ ID NO: 112.

XX

KW Cryptosporidium detection; GP900; P68; cryptopain; cryptosporidiosis.

XX

OS Cryptosporidium parvum.

XX

PN WO200194631-A1.

XX

PD 13-DEC-2001.

XX

PF 14-MAY-2001; 2001WO-US015624.

XX

PR 06-JUN-2000; 2000US-00588995.

XX

PA (REGC ) UNIV CALIFORNIA.

XX

PI Petersen C, Barnes DA, Nelson RG, Gut J;

XX

DR WPI; 2002-566447/60.

XX

PT Detecting Cryptosporidium in biological and environmental samples and  
PT diagnosis of cryptosporidiosis involves, contacting the sample with  
PT Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or RNA.

XX

PS Example 18; Page 154; 157pp; English.

XX

CC The present invention relates to a method of detecting Cryptosporidium in  
CC biological and environmental samples, and of diagnosing  
CC cryptosporidiosis. This involves obtaining a sample and contacting it  
CC with Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or  
CC RNA, or its variant, mutant or fragment. The method is also useful for  
CC detecting and identifying individual Cryptosporidium isolates based on  
CC the genetic characteristics, and for diagnosis of prior or concurrent  
CC Cryptosporidium infection. The present sequence is a C. parvum peptide  
CC sequence used in the exemplification of the invention

XX Sequence 8 AA;

Query Match

Best Local Similarity 71.9%; Score 41; DB 5; Length 8;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WTVNSW 7

DB 2 WTVNSW 8

RESULT 4

AAR77352

ID AAR77352 standard; peptide; 8 AA.

XX

AC AAR77352;

XX

DT 14-MAY-1996 (first entry)

XX

DE Cysteine proteinase derived peptide #3.

XX

KW Primer; amplify; polymerase chain reaction; PCR; Haemonchus contortus;

KW cysteine proteinase; DM.2; DM.3; DM.4; human; DM.4a; DM.5; antigen;

KW vaccine; helminth; parasite; DM.1; ruminant; gastro-intestinal tract.

XX

OS Haemonchus contortus.

XX

PN WO9526402-A1.

XX

PD 05-OCT-1995.



PF 24-MAR-1995; 95WO-GB000665.  
 XX  
 PR 25-MAR-1994; 94GB-00005925.  
 PR 25-MAR-1994; 94GB-00005990.  
 XX  
 PA (MLCW ) MALLINCKRODT VETERINARY INC.  
 XX  
 PI Knox DP, Smith SK, Smith WD, Redmond D, Murray J;  
 XX  
 DR NPI; 1995-351322/45.  
 DR N-PSDB; AAQ94242.  
 XX  
 PT Protective helminth parasite antigen - used in vaccine directed against  
 PT parasitic nematodes of mammalian gastro-intestinal tract e.g. Haemonchus  
 PT contortus.  
 XX  
 PS Example 16; Fig 15; 79pp; English.  
 XX  
 CC The sequences given in AAR77350-53 are peptides derived from the  
 CC canonical haemonchus contortus cysteine proteinase molecule which were  
 CC used in the design of the primers given in AAQ94240-43. These primers  
 CC were used in the cloning of cDNA fragments from the cysteine proteinase  
 CC gene, such as DM.1, DM.2, DM.3, DM.4 and DM.5 (see also AAQ94246-  
 CC 51). The amplified fragments may be expressed in a recombinant cell for  
 CC the production of antigens. These antigens may be used in the preparation  
 CC of a vaccine against helminth parasites in a human or non-human animal  
 XX  
 SQ Sequence 8 AA;  
 Query Match 64.9%; Score 37; DB 2; Length 8;  
 Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 WTVRNSW 7  
 Db 2 WIVKNSW 8  
 RESULT 5  
 AAB76056  
 ID AAB76056 standard; peptide; 9 AA.  
 XX  
 AC AAB76056;  
 XX  
 DT 10-APR-2001 (first entry)  
 XX  
 DE Trypanosoma cruzi HLA-A2 binding peptide.  
 XX  
 KW Human leukocyte antigen; HLA; major histocompatibility complex; MHC;  
 KW cytotoxic T lymphocyte; CTL; human class I MHC; immunogenic;  
 KW HLA binding peptide; immune response; glycoprotein; cytostatic; virucide;  
 KW hepatotropic; antiinflammatory; anti-HIV; vaccine;  
 KW human immunodeficiency virus; protozoacide; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; human papilloma virus; HPV;  
 KW cytomegalovirus; CMV; acquired immunodeficiency syndrome; AIDS;  
 KW renal carcinoma; cervical carcinoma; lymphoma; malaria;  
 KW condyloma acuminatum.  
 XX  
 OS Trypanosoma cruzi.  
 XX  
 PN WO200100225-A1.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 28-JUN-2000; 2000WO-US017842.  
 XX  
 PR 29-JUN-1999; 99US-0141422P.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S;  
 XX  
 DR WPI; 2001-112389/12.

XX  
 PT Composition comprising human leukocyte antigen binding peptide which  
 PT comprises isolated, prepared epitope useful for treating viral infections  
 PT such as acquired immunodeficiency syndrome, and cancer.  
 XX  
 PS Claim 1; Page 46; 58pp; English.  
 XX  
 CC The present invention describes a composition (I) which comprises at  
 CC least one human leukocyte antigen (HLA) binding peptide comprising an  
 CC isolated, prepared epitope comprising one of 547 8-11 residue amino acid  
 CC sequences (S1), given in AAB75803 to AAB76349. (I) has cytostatic,  
 CC virucide, hepatotropic, antiinflammatory, anti-HIV (human  
 CC immunodeficiency virus) and protozoacide activities, which can be used in  
 CC vaccine production and is an inducer of cytotoxic T-cell response. (I) is  
 CC useful for inducing a cytotoxic T cell response against a preselected  
 CC antigen in a patient expressing a specific major histocompatibility  
 CC complex (MHC) class I allele, by contacting cytotoxic T cells (CTLs) from  
 CC the patient with (I). (I) is useful as a vaccine to treat and/or prevent  
 CC viral infection and cancer such as prostate cancer, hepatitis B,  
 CC hepatitis C, human papilloma virus (HPV) infection, cytomegalovirus  
 CC (CMV), acquired immunodeficiency syndrome (AIDS), renal carcinoma,  
 CC cervical carcinoma, lymphoma, malaria, and condyloma acuminatum  
 XX  
 SQ Sequence 9 AA;  
 Query Match 63.2%; Score 36; DB 4; Length 9;  
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 WTVRNSW 7  
 Db 1 WIKNSW 7  
 RESULT 6  
 AAW60399  
 ID AAW60399 standard; peptide; 9 AA.  
 XX  
 AC AAW60399;  
 XX  
 DT 24-AUG-1998 (first entry)  
 XX  
 DE Tumour homing peptide of the invention.  
 XX  
 KW Tumour homing peptide; in vivo panning; murine melanoma; tumour.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9810795-A2.  
 XX  
 PD 19-MAR-1998.  
 XX  
 PF 10-SEP-1997; 97WO-US016086.  
 XX  
 PR 10-SEP-1996; 96US-00710067.  
 XX  
 PA (BURN-) BURNHAM INST.  
 XX  
 PI Ruoslahti E, Pasqualini R;  
 XX  
 DR WPI; 1998-207151/18.  
 XX  
 PT Tumour homing molecules and their conjugates - useful for, e g. directing  
 PT linked moiety to tumour containing angiogenic vasculature.  
 XX  
 PS Example 5; Page 80; 105pp; English.  
 XX  
 CC AAW60390-432 represent peptides recovered from mouse melanomas. The  
 CC peptides are tumour homing peptides, and are identified by in vivo  
 CC panning. The in vivo panning comprises administering a library of diverse  
 CC peptides to a subject having a tumour, collecting a sample of the tumour,  
 CC identifying a peptide that homes to the tumour, collecting a sample of  
 CC normal tissue corresponding to the tumour, and determining that the

CC peptide that homes to the tumour is not present in the normal tissue. The  
 CC tumour homing peptides can be linked to a moiety (e.g. doxorubicin), and  
 CC used to direct the moiety to a tumour

XX Sequence 9 AA;

Query Match 57.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSW 7  
 |||||  
 Db 1 WTCRASW 7

## RESULT 7

AAW93726  
 ID AAW93726 standard; peptide; 9 AA.

XX AC AAW93726;

XX DT 28-JUN-1999 (first entry)

XX DE Mouse B16B15b melanoma derived tumour homing peptide 11.

XX KW Tumour homing; peptide; tumour; diagnosis; endothelial cell; melanoma;  
 KW angiogenic vasculature; anti-tumour; anti-inflammatory; anti-angiogenic;  
 KW anti-arthritis; NGR receptor; inhibitor; angiogenesis; anticancer drug;  
 KW prognosis; inflammation; regeneration; wounded tissue; targeting;  
 KW macular degeneration; diabetic retinopathy; rheumatoid arthritis;  
 KW occlusive thrombus; murine.

XX OS Mus sp.

XX PN WO9913329-A1.

XX PD 18-MAR-1999.

XX PF 08-SEP-1998; 98WO-US018895.

XX PR 10-SEP-1997; 97US-00926914.

XX PR 25-AUG-1998; 98US-00139802.

XX PA (BURN-) BURNHAM INST.

XX PI Ruoslahti E, Pasqualini R;

XX DR WPI; 1999-215158/18.

XX PT Identifying molecules that home to angiogenic vasculature used as targets  
 XX for anticancer agents.

XX PS Example V7; Page 115; 180pp; English.

XX CC This invention describes novel peptides which home to angiogenic  
 CC vasculature, specifically of a tumour and which have anti-tumour, anti-  
 CC inflammatory, anti-angiogenic and anti-arthritis activity. Such molecules  
 CC are identified by treating a purified NGR receptor with a test compound  
 CC and identifying compounds that bind specifically to the NGR receptor. The  
 CC peptides of the invention are inhibitors of angiogenesis and can be used  
 CC to produce conjugates for delivering agents to angiogenic vasculature,  
 CC particularly anticancer drugs or an imaging agent, for diagnosis or  
 CC prognosis. These conjugates may be directed to non-tumour angiogenic  
 CC vasculature, e.g. that present in inflammatory, regenerating or wounded  
 CC tissue, e.g. for treatment of macular degeneration, diabetic retinopathy  
 CC or rheumatoid arthritis. The peptides provide specific targeting to  
 CC tumours, especially their supporting vasculature, since the NGR receptor  
 CC is exposed to the circulation only in angiogenic vasculature. Precise  
 CC targeting should reduce the systemic toxicity of anticancer drugs in the  
 CC conjugates. Complete killing of all target cells may not be essential  
 CC since partial denudation of endothelium may result in an occlusive  
 CC thrombus, and endothelial cells are unlikely to become resistant to  
 CC anticancer agents nor to lose the targeting receptor. AAW93622-W93809 and

CC AAW93843-44 are examples of tumour homing peptides used in the invention  
 XX  
 SQ Sequence 9 AA;

Query Match 57.9%; Score 33; DB 2; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSW 7  
 |||||  
 Db 1 WTCRASW 7

## RESULT 8

AAB21816  
 ID AAB21816 standard; peptide; 9 AA.

XX AC AAB21816;

XX DT 22-MAR-2001 (first entry)

XX DE Murine melanoma homing peptide #11.

XX KW Cytostatic; homing pro-apoptotic conjugate; tumour; antimicrobial;  
 KW breast; prostate; melanoma; cancer; Kaposi's sarcoma; murine.

XX OS Mus sp.

XX PN WO200042973-A2.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-US001602.

XX PR 22-JAN-1999; 99US-00235902.

XX PA (BURN-) BURNHAM INST.

XX PI Ellerby HM, Bredesen DE, Pasqualini R, Ruoslahti EI;

XX DR WPI; 2000-499174/44.

XX PT Homing pro-apoptotic conjugate comprising a tumor homing molecule that  
 XX selectively homes to a mammalian cell type or tissue linked to an  
 XX antimicrobial peptide; useful for the treatment of prostate cancer.

XX PS Example 8; Page 96; 118pp; English.

XX CC The present invention relates to homing pro-apoptotic conjugates,  
 CC comprising of a tumour homing molecule that selectively homes to a  
 CC mammalian cell type or tissue, linked to an antimicrobial peptide. The  
 CC homing pro-apoptotic conjugates are selectively internalised by the  
 CC mammalian cell type or tissue and exhibits high toxicity, especially to  
 CC angiogenic vasculature. The antimicrobial peptide has low mammalian cell  
 CC toxicity when not linked to the tumor homing molecule. The conjugates are  
 CC useful for the treatment of cancer e.g. Kaposi's sarcoma, breast and  
 CC prostate cancer or melanoma. The present sequence is a homing peptide  
 CC isolated in the present invention, which can be conjugated to an  
 CC antimicrobial peptide to make the homing pro-apoptotic conjugates of the  
 CC present invention

XX Sequence 9 AA;

Query Match 57.9%; Score 33; DB 3; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVNSW 7  
 |||||  
 Db 1 WTCRASW 7

## RESULT 9

AAE06394  
 ID AAE06394 standard; peptide; 9 AA.  
 XX  
 AC AAE06394;  
 XX  
 DT 25-SEP-2001 (first entry)  
 XX  
 DE Tumour homing peptide #11 from mouse B16B15b melanoma.  
 XX  
 DE Chimeric prostate-homing pro-apoptotic peptide; prostate-homing peptide;  
 KW antimicrobial peptide; prostate cancer; tumour homing molecule; mouse;  
 KW cytostatic.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200153342-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 XX 16-JAN-2001; 2001WO-US001362.  
 PF  
 XX 21-JAN-2000; 2000US-00489582.  
 PR  
 XX (BURN-) BURNHAM INST.  
 PA  
 XX Ruoslahti EI, Pasqualini R, Arap W, Bredesen DE, Ellerby HM;  
 PI  
 XX WPI; 2001-451901/48.  
 DR  
 XX  
 XX Novel chimeric prostate-homing pro-apoptotic peptide, used to treat  
 PT prostate cancer, comprises a prostate-homing peptide linked to an  
 PT antimicrobial peptide.  
 XX  
 PS Example 8; Page 95; 176pp; English.  
 XX  
 CC The patent discloses novel chimeric prostate-homing pro-apoptotic peptide  
 CC which comprises a prostate-homing peptide linked to an antimicrobial  
 CC peptide, where the chimeric peptide is selectively internalised by and  
 CC exhibits high toxicity to prostate tissue and where the antimicrobial  
 CC peptide has low mammalian cell toxicity when not linked to prostate-  
 CC homing peptide. The chimeric peptide is used to direct an antimicrobial  
 CC peptide in vivo to a prostate cancer, to induce selective toxicity in  
 CC vivo in a prostate cancer, and to treat a patient with prostate cancer.  
 CC The present sequence is a tumour homing peptide from mouse B16B15b  
 CC melanoma. This sequence is useful in the homing of pro-apoptotic  
 CC conjugates of the invention  
 XX  
 SQ Sequence 9 AA;  
 Query Match 57.9%; Score 33; DB 4; Length 9;  
 Best Local Similarity 71.4%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 WTVRNSW 7  
 Db 1 WTCRASW 7  
 RESULT 10  
 AAB85132  
 ID AAB85132 standard; peptide; 5 AA.  
 XX  
 AC AAB85132;  
 XX  
 DT 22-AUG-2001 (first entry)  
 XX  
 XX C. parvum cryptopain protein fragment.  
 DE  
 XX Cryptosporidium; infection; cryptopain; antiparasitic; vaccine;  
 KW immunization.  
 KW  
 XX Cryptosporidium parvum.  
 OS  
 XX

PN US6254869-B1.  
 XX  
 PD 03-JUL-2001.  
 XX  
 PF 27-MAR-1997; 97US-00827171.  
 XX  
 PR 27-MAR-1996; 96US-0014233P.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Petersen C, Huang J;  
 XX  
 DR WPI; 2001-388923/41.  
 XX  
 XX New vaccines and cryptopain protein for diagnosis and treatment of  
 PT Cryptosporidium species infection.  
 XX  
 PS Example 2; Col 19; 32pp; English.  
 XX  
 CC The invention relates to vaccines and cryptopain protein for diagnosis  
 CC and treatment of Cryptosporidium species infection. Cryptopain protein  
 CC can be expressed by standard recombinant methodology. The vaccines  
 CC comprising the cryptopain protein or its fragments are useful for active  
 CC immunization of animals and humans against Cryptosporidium infection, or  
 CC for production of passive immune products in admixture with an adjuvant.  
 CC Cryptopain can be used for prophylactic, therapeutic, diagnostic and  
 CC detection purposes. The present sequence represents a peptide fragment of  
 CC the C. parvum cryptopain protein  
 XX  
 SQ Sequence 5 AA;  
 Query Match 52.6%; Score 30; DB 4; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 VRNSW 7  
 Db 1 VRNSW 5  
 RESULT 11  
 ABU04066  
 ID ABU04066 standard; peptide; 5 AA.  
 XX  
 AC ABU04066;  
 XX  
 DT 27-SEP-2002 (first entry)  
 XX  
 DE C parvum cryptopain peptide fragment SEQ ID NO: 113.  
 XX  
 KW Cryptosporidium detection; GP900; P68; cryptopain; cryptosporidiosis.  
 XX  
 OS Cryptosporidium parvum.  
 XX  
 PN WO200194631-A1.  
 PD 13-DEC-2001.  
 XX  
 XX 14-MAY-2001; 2001WO-US015624.  
 PF  
 XX 06-JUN-2000; 2000US-00588995.  
 PR  
 XX (REGC ) UNIV CALIFORNIA.  
 PA  
 XX Petersen C, Barnes DA, Nelson RG, Gut J;  
 PI  
 XX WPI; 2002-566447/60.  
 DR  
 XX Detecting Cryptosporidium in biological and environmental samples and  
 PT diagnosis of cryptosporidiosis involves, contacting the sample with  
 PT Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or RNA.  
 XX  
 XX Disclosure; Page 154; 157pp; English.  
 PS

XX The present invention relates to a method of detecting Cryptosporidium in  
 CC biological and environmental samples, and of diagnosing  
 CC cryptosporidiosis. This involves obtaining a sample and contacting it  
 CC with Cryptosporidium GP900, P68 or cryptopain antigen, antibody, DNA or  
 CC RNA, or its variant, mutant or fragment. The method is also useful for  
 CC detecting and identifying individual Cryptosporidium isolates based on  
 CC the genetic characteristics, and for diagnosis of prior or concurrent  
 CC Cryptosporidium infection. The present sequence is a C. parvum peptide  
 CC sequence used in the exemplification of the invention  
 XX  
 XX Sequence 5 AA;

Query Match 52.6%; Score 30; DB 5; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRNSW 7  
 Db 1 VRNSW 5

RESULT 12  
 ADQ95813  
 ID ADQ95813 standard; peptide; 5 AA.

XX AC  
 XX ADQ95813;  
 XX  
 DT 23-SEP-2004 (first entry)

XX DE Cryptosporidium parvum cryptopain fragment #3.

XX KW Cryptosporidium infection; antigen; cryptopain;  
 XX cathepsin L-like cysteine proteinase; passive immunity; vaccine; enzyme.  
 XX OS  
 XX Cryptosporidium parvum.

XX PN US6759044-B1.  
 XX PD 06-JUL-2004.

XX PF 20-JUN-2000; 2000US-00598062.

XX PR 27-MAR-1996; 96US-0014233P.

XX PR 27-MAR-1997; 97US-00827171.

XX FA (REGC) UNIV CALIFORNIA.

XX PI Petersen C, Huang J;

XX DR WPI; 2004-515391/49.

XX PT New cryptopain antibodies, useful for treating or inhibiting  
 CC Cryptosporidium infection, or as vaccine for to provide passive immunity  
 CC to Cryptosporidium infection.

XX PS Example 2; SEQ ID NO 15; 33pp; English.

XX CC The invention relates to an antibody produced against and binding to  
 CC Cryptosporidium antigen, cryptopain. The invention also provides  
 CC polynucleotides encoding cryptopain which is a cathepsin L-like cysteine  
 CC proteinase. The antibody is useful for detecting and treating  
 CC Cryptosporidium infection, for providing passive immunity or for  
 CC inhibiting existing Cryptosporidium infection. The present sequence is  
 CC Cryptosporidium parvum cryptopain active site fragment comprising  
 CC conserved Asn residue.

XX Sequence 5 AA;

Query Match 52.6%; Score 30; DB 8; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VRNSW 7  
 Db 1 VRNSW 5

RESULT 13

ADH12805

ID ADH12805 standard; peptide; 7 AA.

XX AC ADH12805;

XX DT 11-MAR-2004 (first entry)

XX DE Abalone cellulase N-terminal peptide, SEQ ID NO:6.

XX KW Abalone; cellulase; liver pancreas; de-inking; decolouring;  
 KW paper manufacture; detergent; food manufacture; animal fodder;  
 KW biomass degradation; energy production; cello-oligosaccharide production;  
 KW cellobiose; cellotriose; waste recycling; EC 3.2.1.4.

XX OS Halliotis discus; subsp. hannai.

XX PN JP2003235552-A.

XX PD 26-AUG-2003.

XX PF 13-FEB-2002; 2002JP-00034852.

XX PR 13-FEB-2002; 2002JP-00034852.

XX PA (HOKK-) HOKKAIDO TLO KK.

XX DR WPI; 2004-147477/15.

XX PT Novel cellulase originating in spiral shells capable of degrading  
 CC cellulose, useful for producing cell oligosaccharide such as cellobiose  
 CC and a cellitriose.

XX PS Example 3; SEQ ID NO 6; 21pp; Japanese.

XX CC The invention relates to a cellulase (ADH12819) from the abalone Halliotis  
 CC discus hannai. The cellulase is present in high levels in the liver  
 CC pancreas of the abalone, particularly the anterior portion. The mature  
 CC cellulase (EC 3.2.1.4) has an N-terminal sequence given in ADH12801, and  
 CC has a molecular weight of 66 kD, a pH optimum of 5.5-8.0, an optimum  
 CC temperature of 35-40 degrees Celsius, and is stable at 40 degrees Celsius  
 CC or below. The invention also encompasses a method for the preparation of  
 CC the cellulase from abalone internal organs, and further discloses a cDNA  
 CC sequence (ADH12800) encoding the cellulase. The abalone cellulase is  
 CC useful in the de-inking and decolouring processes in paper manufacturing,  
 CC in detergent compositions, and in the manufacture of foods or animal  
 CC fodder. It is also useful in the degradation of biomass for energy  
 CC production and for the production of cello-oligosaccharides such as  
 CC cellobiose and cellotriose. The cellulase of the invention is extracted  
 CC from parts of abalone which are inedible, thousands of tonnes of which  
 CC are discarded during processing of the shellfish as a foodstuff. The  
 CC method of the invention permits helps to reduce the environmental impact  
 CC of this waste, and permits the inexpensive and rapid production of large  
 CC quantities of cellulase. The present sequence represents an N-terminal  
 CC peptide fragment of abalone cellulase used to design a degenerate PCR  
 CC primer in an example of the invention.

XX Sequence 7 AA;

Query Match 50.9%; Score 29; DB 8; Length 7;  
 Best Local Similarity 57.1%; Pred. No. 1.7e+06;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TVRNSWD 8  
 Db 1 TISHNWD 7

## RESULT 14

AB85131  
ID AAB85131 standard; peptide; 8 AA.

XX  
AC AAB85131;  
XX  
DT 22-AUG-2001 (first entry)  
XX  
DE P. vinckei cysteine proteinase conserved motif.

XX  
KW Cryptosporidium; infection; cryptopain; antiparasitic; vaccine;  
immunization; cysteine proteinase.

XX  
OS Plasmodium vinckei.

XX  
FH Key Location/Qualifiers

FT Misc-difference 4 /label= Val or Ile

FT Misc-difference 5 /label= Lys or Arg

XX  
PN US6254869-B1.

XX  
PD 03-JUL-2001.

XX  
PF 27-MAR-1997; 97US-00827171.

XX  
PR 27-MAR-1996; 96US-0014233P.

XX  
PA (REGC ) UNIV CALIFORNIA.

XX  
PI Petersen C, Huang J;

XX  
DR WPI; 2001-388923/41.

XX  
PT New vaccines and cryptopain protein for diagnosis and treatment of  
PT Cryptosporidium species infection.

XX  
PS Example 2; Col 19; 32pp; English.

XX  
CC The invention relates to vaccines and cryptopain protein for diagnosis  
and treatment of Cryptosporidium species infection. Cryptopain protein  
can be expressed by standard recombinant methodology. The vaccines  
comprising the cryptopain protein or its fragments are useful for active  
immunization of animals and humans against Cryptosporidium infection, or  
for production of passive immune products in admixture with an adjuvant.  
Cryptopain can be used for prophylactic, therapeutic, diagnostic and  
detection purposes. The present sequence represents a conserved motif  
from P. vinckei cysteine proteinase, used for designing degenerate  
primers for amplifying DNA fragments of C. parvum cryptopain gene

XX  
SQ Sequence 8 AA;

Query Match 50.9%; Score 29; DB 4; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.7e+06;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WTVRNSW 7

Db 2 WIXXNSW 8

## RESULT 15

ADQ95812

ID ADQ95812 standard; peptide; 8 AA.

XX  
AC ADQ95812;

XX  
DT 23-SEP-2004 (first entry)

XX  
DE Cryptosporidium parvum cryptopain fragment #2.

XX  
KW Cryptosporidium infection; antigen; cryptopain;

KW  
XX cathepsin L-like cysteine proteinase; passive immunity; vaccine; enzyme.  
XX  
OS Cryptosporidium parvum.

XX  
FH Key Location/Qualifiers

FT Misc-difference 4 /label= Val, Ile

FT Misc-difference 5 /label= Lys, Arg

XX  
PN US6759044-B1.

XX  
PD 06-JUL-2004.

XX  
PF 20-JUN-2000; 2000US-00598062.

XX  
PR 27-MAR-1996; 96US-0014233P.

XX  
PR 27-MAR-1997; 97US-00827171.

XX  
PA (REGC ) UNIV CALIFORNIA.

XX  
PI Petersen C, Huang J;

XX  
DR WPI; 2004-515391/49.

XX  
PT New cryptopain antibodies, useful for treating or inhibiting  
PT Cryptosporidium infection, or as vaccine for to provide passive immunity  
to Cryptosporidium infection.

XX  
PS Example 2; SEQ ID NO 14; 33pp; English.

XX  
CC The invention relates to an antibody produced against and binding to  
Cryptosporidium antigen, cryptopain. The invention also provides  
polynucleotides encoding cryptopain which is a cathepsin L-like cysteine  
proteinase. The antibody is useful for detecting and treating  
Cryptosporidium infection, for providing passive immunity or for  
inhibiting existing Cryptosporidium infection. The present sequence is  
Cryptosporidium parvum cryptopain fragment.

XX  
SQ Sequence 8 AA;

Query Match 50.9%; Score 29; DB 8; Length 8;  
Best Local Similarity 57.1%; Pred. No. 1.7e+06;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 WTVRNSW 7

Db 2 WIXXNSW 8

Search completed: October 27, 2004, 18:50:15  
Job time : 153 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 18:50:23 ; Search time 126 Seconds  
(without alignments)  
23.158 Million cell updates/sec

Title: US-09-867-159A-5  
Perfect score: 57  
Sequence: 1 WTVRNSWDT 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 132014

Minimum DB seq length: 0  
Maximum DB seq length: 9

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	9	10	US-09-867-159A-5
2	33	57.9	9	9	US-09-765-086-116
3	33	57.9	9	14	US-10-264-374-116
4	33	57.9	9	14	US-10-375-992-116
5	33	57.9	9	15	US-10-264-374-116
6	33	57.9	9	16	US-10-375-992-116
7	27	47.4	9	9	US-09-834-765-155
8	26	45.6	7	16	US-10-363-204-24
9	26	45.6	8	14	US-10-190-082-286
10	26	45.6	9	14	US-10-024-652-130
11	26	45.6	9	14	US-10-024-652-981
12	25	43.9	9	10	US-09-793-451-30
13	25	43.9	9	10	US-09-793-451-235

14	25	43.9	9	10	US-09-793-451-353	Sequence 353, App
15	25	43.9	9	10	US-09-793-451-649	Sequence 649, App
16	25	43.9	9	10	US-09-942-052-119	Sequence 119, App
17	25	43.9	9	14	US-10-283-722-30	Sequence 30, Appl
18	25	43.9	9	14	US-10-283-722-235	Sequence 235, App
19	25	43.9	9	14	US-10-283-722-353	Sequence 353, App
20	25	43.9	9	14	US-10-283-722-649	Sequence 649, App
21	25	43.9	9	14	US-10-283-903-30	Sequence 30, Appl
22	25	43.9	9	14	US-10-283-903-235	Sequence 235, App
23	25	43.9	9	14	US-10-283-903-353	Sequence 353, App
24	25	43.9	9	14	US-10-283-903-649	Sequence 649, App
25	24	42.1	7	14	US-10-190-082-16	Sequence 16, Appl
26	24	42.1	8	14	US-10-351-641-1643	Sequence 1643, App
27	24	42.1	9	14	US-10-334-726-184	Sequence 184, App
28	24	42.1	9	14	US-10-334-726-295	Sequence 295, App
29	24	42.1	9	15	US-10-609-217-664	Sequence 664, App
30	24	42.1	9	15	US-10-632-388-664	Sequence 664, App
31	24	42.1	9	15	US-10-651-723-664	Sequence 664, App
32	24	42.1	9	15	US-10-645-761-664	Sequence 664, App
33	24	42.1	9	15	US-10-666-696-664	Sequence 664, App
34	24	42.1	9	15	US-10-653-048-664	Sequence 664, App
35	23	40.4	5	16	US-10-705-195-17	Sequence 17, Appl
36	23	40.4	5	16	US-10-714-564A-471	Sequence 471, App
37	23	40.4	6	14	US-10-072-602B-626	Sequence 626, App
38	23	40.4	6	14	US-10-072-602B-628	Sequence 628, App
39	23	40.4	7	14	US-10-072-602B-627	Sequence 627, App
40	23	40.4	7	14	US-10-190-082-129	Sequence 129, App
41	23	40.4	7	14	US-10-190-082-133	Sequence 133, App
42	23	40.4	8	14	US-10-347-562-22	Sequence 22, Appl
43	23	40.4	8	14	US-10-351-641-1562	Sequence 1562, App
44	23	40.4	8	14	US-10-351-641-1638	Sequence 1638, App
45	23	40.4	8	14	US-10-351-641-1639	Sequence 1639, App

## ALIGNMENTS

RESULT 1  
US-09-867-159A-5  
; Sequence 5, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANITALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; TITLE OF INVENTION: and at least one anti-histamine compound  
; FILE REFERENCE: BL12812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(9)  
; OTHER INFORMATION: Comprises epitope from cystine protease.  
US-09-867-159A-5

Query Match 100.0%; Score 57; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 WTVRNSWDT 9  
Db 1 WTVRNSWDT 9

```
RESULT 2
US-09-765-086-116
; Sequence 116, Application US/09765086
; Patent No. US20010046498A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Wadhi, Arad
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Ellerby, H. Michael
; TITLE OF INVENTION: Chimeric Prostate-Homing Peptides With
; TITLE OF INVENTION: Pro-Apoptotic Activity
; FILE REFERENCE: P-LJ 3844
; CURRENT FILING DATE: 2001-01-17
; CURRENT APPLICATION NUMBER: US/09/765,086
; PRIOR APPLICATION NUMBER: US 09/499,582
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 235
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-09-765-086-116
Query Match 57.9%; Score 33; DB 9; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
DB 1 WTCRASW 7

RESULT 3
US-10-264-374-116
; Sequence 116, Application US/10264374
; Publication No. US20030113320A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; FILE REFERENCE: P-LJ 3203
; CURRENT FILING DATE: 2002-10-03
; CURRENT APPLICATION NUMBER: US/10/264,374
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 08/926,914
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 08/710,067
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-264-374-116
Query Match 57.9%; Score 33; DB 14; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
DB 1 WTCRASW 7

RESULT 4
US-10-375-992-116
; Sequence 116, Application US/10375992
; Publication No. US20030152578A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Tumor Homing Molecules, Conjugates
; TITLE OF INVENTION: Derived Therefrom, and Methods of Using Same
; NUMBER OF SEQUENCES: 199
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/375,992
; FILING DATE: 27-Feb-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,914
; FILING DATE: 10-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-10-375-992-116
Query Match 57.9%; Score 33; DB 14; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WTVRNSW 7
DB 1 WTCRASW 7

RESULT 5
US-10-264-374-116
; Sequence 116, Application US/10264374
; Publication No. US20040096441A9
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: NGR Receptor and Methods of Identifying Tumor Homing
; TITLE OF INVENTION: Molecules That Home to Angiogenic Vasculature Using
; FILE REFERENCE: P-LJ 3203
; CURRENT FILING DATE: 2002-10-03
; CURRENT APPLICATION NUMBER: US/10/264,374
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 08/926,914
; PRIOR FILING DATE: 1997-09-10
```



```

; PRIOR APPLICATION NUMBER: 08/710,067
; PRIOR FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-264-374-116

Query Match          57.9%; Score 33; DB 15; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
      |||||
Db      1 WTCRASW 7

RESULT 6
US-10-375-992-116
; Sequence 116, Application US/10375992
; Publication No. US20040131623A9
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; TITLE OF INVENTION: Tumor Homing Molecules, Conjugates
; NUMBER OF SEQUENCES: 199
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cambell & Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/375,992
; FILING DATE: 27-Feb-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/926,914
; FILING DATE: 10-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Cambell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2725
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-10-375-992-116

Query Match          57.9%; Score 33; DB 15; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
      |||||
Db      1 WTCRASW 7

US-09-834-765-155
; Sequence 155, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GIP-BINDING PROTEIN USEFUL IN TREATMENT
; FILE REFERENCE: 129.6USUI
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 155
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-155

Query Match          47.4%; Score 27; DB 9; Length 9;
Best Local Similarity 62.5%; Pred. No. 1.2e+06;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 TVRNSWDT 9
      |||||
Db      2 TVQLQMDT 9

RESULT 8
US-10-363-204-24
; Sequence 24, Application US/10363204
; Publication No. US20040170955A1
; GENERAL INFORMATION:
; APPLICANT: Board of Regents, The University of Texas System
; TITLE OF INVENTION: Human and Mouse Targeting Peptides Identified by Phage Display
; FILE REFERENCE: 005774.P003PCT
; CURRENT APPLICATION NUMBER: US/10/363,204
; CURRENT FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: (1)..(7)
; OTHER INFORMATION: synthetic construct
US-10-363-204-24

Query Match          45.6%; Score 26; DB 16; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.2e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 WTVRN 5
      |||||
Db      2 WTVRD 6

RESULT 9
US-10-190-082-286
; Sequence 286, Application US/10190082
; Publication No. US20030148264A1
```

; GENERAL INFORMATION:  
 ; APPLICANT: Lasky, Lawrence A.  
 ; APPLICANT: Sidhu, Sachdev S.  
 ; APPLICANT: Held, Heike A.  
 ; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS  
 ; FILE REFERENCE: P190SR1  
 ; CURRENT APPLICATION NUMBER: US/10/190,082  
 ; CURRENT FILING DATE: 2002-07-03  
 ; PRIOR APPLICATION NUMBER: US 60/303,634  
 ; PRIOR FILING DATE: 2001-07-06  
 ; NUMBER OF SEQ ID NOS: 683  
 ; SEQ ID NO 286  
 ; LENGTH: 8  
 ; TYPE: PRT  
 ; ORGANISM: Artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic  
 US-10-190-082-286

Query Match 45.6%; Score 26; DB 14; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 TVRNSW 7  
 ||| :||  
 DB 2 TVRETW 7

RESULT 10  
 US-10-024-652-130  
 ; Sequence 130, Application US/10024652  
 ; Publication No. US20030219738A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Agensys, Inc  
 ; APPLICANT: Challita-Eid, Pia M.  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Afar, Daniel E.H.  
 ; APPLICANT: Hubert, Rene S.  
 ; APPLICANT: Mitchell, Steve Chappell  
 ; APPLICANT: Levin, Eiana  
 ; APPLICANT: Morrison, Karen Jane Meyrick  
 ; APPLICANT: Raitano, Arthur B.  
 ; APPLICANT: Jakobovits, Aya  
 ; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
 ; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
 ; TITLE OF INVENTION: Detection of Cancer  
 ; FILE REFERENCE: 51158-20025.00  
 ; CURRENT APPLICATION NUMBER: US/10/024,652  
 ; CURRENT FILING DATE: 2002-06-28  
 ; PRIOR APPLICATION NUMBER: 60/256,210  
 ; PRIOR FILING DATE: 2000-12-15  
 ; NUMBER OF SEQ ID NOS: 2598  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 130  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: homo sapien  
 US-10-024-652-130

Query Match 45.6%; Score 26; DB 14; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+06;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TVRNSWDT 9  
 ||| :|||  
 DB 2 TFRINDT 9

RESULT 11  
 US-10-024-652-981  
 ; Sequence 981, Application US/10024652  
 ; Publication No. US20030219738A1  
 ; GENERAL INFORMATION:

; APPLICANT: Agensys, Inc  
 ; APPLICANT: Challita-Eid, Pia M.  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Afar, Daniel E.H.  
 ; APPLICANT: Hubert, Rene S.  
 ; APPLICANT: Mitchell, Steve Chappell  
 ; APPLICANT: Levin, Eiana  
 ; APPLICANT: Morrison, Karen Jane Meyrick  
 ; APPLICANT: Raitano, Arthur B.  
 ; APPLICANT: Jakobovits, Aya  
 ; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc  
 ; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and  
 ; TITLE OF INVENTION: Detection of Cancer  
 ; FILE REFERENCE: 51158-20025.00  
 ; CURRENT APPLICATION NUMBER: US/10/024,652  
 ; CURRENT FILING DATE: 2002-06-28  
 ; PRIOR APPLICATION NUMBER: 60/256,210  
 ; PRIOR FILING DATE: 2000-12-15  
 ; NUMBER OF SEQ ID NOS: 2598  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 981  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: homo sapien  
 US-10-024-652-981

Query Match 45.6%; Score 26; DB 14; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.2e+06;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 TVRNSWDT 9  
 ||| :|||  
 DB 2 TFRINDT 9

RESULT 12  
 US-09-793-451-30  
 ; Sequence 30, Application US/09793451  
 ; Publication No. US20030157597A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Gazelle S. Rastegar  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Rene S. Hubert  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY  
 ; TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS  
 ; FILE REFERENCE: 129.2USU2  
 ; CURRENT APPLICATION NUMBER: US/09/793,451  
 ; CURRENT FILING DATE: 2001-02-26  
 ; PRIOR APPLICATION NUMBER: 60/184,558  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: 60/218,856  
 ; PRIOR FILING DATE: 2000-07-13  
 ; NUMBER OF SEQ ID NOS: 752  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 30  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-793-451-30

Query Match 43.9%; Score 25; DB 10; Length 9;  
 Best Local Similarity 42.9%; Pred. No. 1.2e+06;  
 Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WTVRNSW 7  
 ||| :||  
 DB 1 WTYSQW 7

```
RESULT 13
US-09-793-451-235
; Sequence 235, Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 235
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-235

Query Match      43.9%; Score 25; DB 10; Length 9;
Best Local Similarity 42.9%; Pred. No. 1.2e+06;
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
      |||
Db      1 WTYSQGW 7

RESULT 14
US-09-793-451-353
; Sequence 353, Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-353

Query Match      43.9%; Score 25; DB 10; Length 9;
Best Local Similarity 42.9%; Pred. No. 1.2e+06;
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
      |||
Db      1 WTYSQGW 7

RESULT 15
US-09-793-451-649
; Sequence 649, Application US/09793451
; Publication No. US20030157597A1
; GENERAL INFORMATION:
; APPLICANT: Arthur B. Raitano
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Gazelle S. Rastegar
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Rene S. Hubert
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
; FILE OF INVENTION: EXPRESSED IN VARIOUS CANCERS
; FILE REFERENCE: 129.2USU2
; CURRENT APPLICATION NUMBER: US/09/793,451
; CURRENT FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/184,558
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/218,856
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 649
; LENGTH: 9
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-793-451-649

Query Match      43.9%; Score 25; DB 10; Length 9;
Best Local Similarity 42.9%; Pred. No. 1.2e+06;
Matches 3; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 WTVRNSW 7
      |||
Db      1 WTYSQGW 7

Search completed: October 27, 2004, 18:57:36
Job time : 131 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 20:26:03 ; Search time 47.3581 Seconds  
(without alignments)  
8987.005 Million cell updates/sec

Title: US-09-867-159A-6

Perfect score: 9

Sequence: 1 ggggggggg 9

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_hgt:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	100.0	9	6	AX586977 Sequence
2	9	100.0	11	6	CQ779064 Sequence
3	9	100.0	11	6	CQ836917 Sequence
4	9	100.0	11	6	CQ836952 Sequence
5	9	100.0	11	6	AX629036 Sequence
6	9	100.0	12	6	AR066414 Sequence
7	9	100.0	12	6	AR094606 Sequence
8	9	100.0	12	6	AR242058 Sequence
9	9	100.0	12	6	AX283723 Sequence
10	9	100.0	12	6	AX586978 Sequence
11	9	100.0	14	6	AR403419 Sequence
12	9	100.0	14	6	AR403420 Sequence
13	9	100.0	14	6	BD068919 Enzymatic
14	9	100.0	14	6	BD068920 Enzymatic
15	9	100.0	15	6	A88671 Sequence 81
16	9	100.0	15	6	A90638 Sequence 81
17	9	100.0	15	6	AR084532 Sequence
18	9	100.0	15	6	BD247095 Improved
19	9	100.0	15	6	AR278935 Sequence

20	9	100.0	15	6	AX027102 Sequence
21	9	100.0	15	6	BD066184 An antise
22	9	100.0	16	6	A88144 Sequence 29
23	9	100.0	16	6	A90111 Sequence 29
24	9	100.0	16	6	AR328248 Sequence
25	9	100.0	16	6	AR435743 Sequence
26	9	100.0	16	6	BD065657 An antise
27	9	100.0	17	6	A27313 Synthetic b
28	9	100.0	17	6	A27314 Synthetic b
29	9	100.0	17	6	A88670 Sequence 81
30	9	100.0	17	6	A90637 Sequence 81
31	9	100.0	17	6	AR069862 Sequence
32	9	100.0	17	6	AR164080 Sequence
33	9	100.0	17	6	AR164081 Sequence
34	9	100.0	17	6	AR242713 Sequence
35	9	100.0	17	6	AR242714 Sequence
36	9	100.0	17	6	AR381868 Sequence
37	9	100.0	17	6	AR381869 Sequence
38	9	100.0	17	6	AX216345 Sequence
39	9	100.0	17	6	AX216346 Sequence
40	9	100.0	17	6	AX216347 Sequence
41	9	100.0	17	6	AX216894 Sequence
42	9	100.0	17	6	AX216895 Sequence
43	9	100.0	17	6	AX216896 Sequence
44	9	100.0	17	6	AX475351 Sequence
45	9	100.0	17	6	AX475352 Sequence

#### ALIGNMENTS

RESULT 1  
AX586977  
LOCUS AX586977 9 bp DNA linear PAT 10-JAN-2003  
DEFINITION Sequence 6 from Patent WO2078736.  
ACCESSION AX586977  
VERSION AX586977.1 GI:27655852  
KEYWORDS  
SOURCE Dermatophagoides pteronyssinus (European house dust mite)  
ORGANISM Dermatophagoides pteronyssinus  
REFERENCE 1  
AUTHORS Loria, E., Terrase, G. and Trehin, Y.  
TITLE Antiallergic pharmaceutical composition  
JOURNAL Patent: WO 02078736-A 6 10-OCT-2002;  
Antiallis (FR)  
FEATURES  
source location/Qualifiers  
1..9  
/organism="Dermatophagoides pteronyssinus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:6956"

ORIGIN  
Query Match 100.0%; Score 9; DB 6; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.2e+09;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCGGCG 9  
Db 1 GCGGCGGCG 9  
RESULT 2  
CQ779064/c  
LOCUS CQ779064 11 bp DNA linear PAT 11-MAR-2004  
DEFINITION Sequence 8 from Patent WO2004015099.  
ACCESSION CQ779064  
VERSION CQ779064.1 GI:45381711  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct

artificial sequences.

REFERENCE 1.  
 AUTHORS Biemann, R., Denoel, P., Feron, C., Goraj, K., Jennings, M. P., Poolman, J. and Weynants, V.  
 TITLE Vaccine composition  
 JOURNAL Patent: WO 2004015099-A 8 19-FEB-2004;  
 GlaxoSmithKline Biologicals S.A. (BE); THE UNIVERSITY OF QUEENSLAND (AU)

FEATURES  
 source Location/Qualifiers  
 1. .11  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="primer"

ORIGIN

Query Match 100.0%; Score 9; DB 6; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 9 GCGGCGGCG 1

RESULT 3  
 LOCUS CQ836917/c 11 bp DNA linear PAT 29-JUL-2004  
 DEFINITION Sequence 1975 from Patent WO2004059001.  
 ACCESSION CQ836917  
 VERSION CQ836917.1 GI:50836451  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 Petersohn, D., Schlotmann, K., Gassenmeier, T., Holtkoetter, O.,  
 Conradt, M. and Hofmann, K.  
 TITLE Method for determining markers of human facial skin  
 JOURNAL Patent: WO 2004059001-A 1975 15-JUL-2004;  
 Henkel Kommanditgesellschaft auf Aktien (DE)

FEATURES  
 source Location/Qualifiers  
 1. .11  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 9; DB 6; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 11 GCGGCGGCG 3

RESULT 4  
 LOCUS CQ836952 11 bp DNA linear PAT 29-JUL-2004  
 DEFINITION Sequence 2010 from Patent WO2004059001.  
 ACCESSION CQ836952  
 VERSION CQ836952.1 GI:50836486  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 Petersohn, D., Schlotmann, K., Gassenmeier, T., Holtkoetter, O.,  
 Conradt, M. and Hofmann, K.  
 TITLE Method for determining markers of human facial skin

JOURNAL Patent: WO 2004059001-A 2010 15-JUL-2004;  
 Henkel Kommanditgesellschaft auf Aktien (DE)

FEATURES  
 source Location/Qualifiers  
 1. .11  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 9; DB 6; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 1 GCGGCGGCG 9

RESULT 5  
 LOCUS AX629036 11 bp DNA linear PAT 21-FEB-2003  
 DEFINITION Sequence 6077 from Patent WO2003774.  
 ACCESSION AX629036  
 VERSION AX629036.1 GI:28457074  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 Petersohn, D., Conradt, M. and Hofmann, K.  
 TITLE Method for determining homeostasis of the skin  
 JOURNAL Patent: WO 02053774-A 6077 11-JUL-2002;  
 Henkel Kommanditgesellschaft auf Aktien (DE)

FEATURES  
 source Location/Qualifiers  
 1. .11  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 9; DB 6; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 1 GCGGCGGCG 9

RESULT 6  
 LOCUS AR066414/c 12 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 38 from patent US 5849995.  
 ACCESSION AR066414  
 VERSION AR066414.1 GI:5996630  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 12)  
 AUTHORS Hayden, M., Lin, B. and Nasir, J.  
 TITLE Mouse model for Huntington's Disease and related DNA sequences  
 JOURNAL Patent: US 5849995-A 38 15-DEC-1998;  
 Location/Qualifiers  
 1. .12  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 9; DB 6; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

artificial sequences.

REFERENCE 1  
AUTHORS Yu, H., Turner, J. G. and Tan, J.  
TITLE Regulation of systemic immune responses utilizing soluble cd40  
JOURNAL Patent: WO 0178769-A 1 25-OCT-2001;  
UNIVERSITY OF SOUTH FLORIDA (US)  
FEATURES Location/Qualifiers  
source 1..12  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="synthetic construct"

ORIGIN  
Query Match 100.0%; Score 9; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
| | | | | | | | | |  
Db 11 GCGGCGGCG 3  
| | | | | | | | | |

RESULT 10  
AX586978  
LOCUS AX586978 12 bp DNA linear PAT 10-JAN-2003  
DEFINITION Sequence 7 from Patent WO02078736.  
ACCESSION AX586978  
VERSION AX586978.1 GI:27655853  
KEYWORDS  
SOURCE Dermatophagoides pteronyssinus (European house dust mite)  
ORGANISM Dermatophagoides pteronyssinus  
REFERENCE 1  
AUTHORS Loria, E., Terras, G. and Trehan, Y.  
TITLE Antiallergic pharmaceutical composition  
JOURNAL Patent: WO 02078736-A 7 10-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..12  
/organism="Dermatophagoides pteronyssinus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6956"

ORIGIN  
Query Match 100.0%; Score 9; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
| | | | | | | | | |  
Db 4 GCGGCGGCG 12  
| | | | | | | | | |

RESULT 11  
AR403419/c  
LOCUS AR403419 14 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 1759 from patent US 6623962.  
ACCESSION AR403419  
VERSION AR403419.1 GI:40150869  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Akhtar, S., Fell, P. and McSwiggen, J. A.  
TITLE Enzymatic nucleic acid treatment of diseases of conditions related to levels of epidermal growth factor receptors  
JOURNAL Patent: US 6623962-A 1759 23-SEP-2003;  
FEATURES Location/Qualifiers  
source 1..14

Qy 1 GCGGCGGCG 9  
| | | | | | | | | |  
Db 11 GCGGCGGCG 3  
| | | | | | | | | |

RESULT 7  
AR084606  
LOCUS AR084606 12 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 95 from patent US 5981185.  
ACCESSION AR084606  
VERSION AR084606.1 GI:10011377  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 12)  
AUTHORS Matson, R. S., Coassin, P. J., Rampal, J. B. and Caskey, C. Thomas.  
TITLE Oligonucleotide repeat arrays  
JOURNAL Patent: US 5981185-A 95 09-NOV-1999;  
FEATURES Location/Qualifiers  
source 1..12  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 100.0%; Score 9; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
| | | | | | | | | |  
Db 1 GCGGCGGCG 9  
| | | | | | | | | |

RESULT 8  
AR242058  
LOCUS AR242058 12 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 346 from patent US 6472154.  
ACCESSION AR242058  
VERSION AR242058.1 GI:27287870  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 12)  
AUTHORS Garner, H. R., Wren, J. D., Minna, J. D. and Fondon, J. W. III.  
TITLE Polymorphic repeats in human genes  
JOURNAL Patent: US 6472154-A 346 29-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..12  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 9; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
| | | | | | | | | |  
Db 1 GCGGCGGCG 9  
| | | | | | | | | |

RESULT 9  
AX283723/c  
LOCUS AX283723 12 bp DNA linear PAT 20-NOV-2001  
DEFINITION Sequence 1 from Patent WO0178769.  
ACCESSION AX283723  
VERSION AX283723.1 GI:17044434  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct

```

/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 9; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
    |||||
Db 13 GCGGCGGCG 5

RESULT 12
AR403420/c
LOCUS      AR403420      14 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 1760 from patent US 6623962.
ACCESSION  AR403420
VERSION     AR403420.1 GI:40150870
KEYWORDS
ORGANISM    Unknown.
SOURCE      Unknown.
REFERENCE   1 (bases 1 to 14)
AUTHORS    Akhtar,S., Fell,P. and McSwiggen,J.A.
TITLE      Enzymatic nucleic acid treatment of diseases or conditions related
           to levels of epidermal growth factor receptors
JOURNAL    Patent: US 6623962-A 1760 23-SEP-2003;
FEATURES    Location/Qualifiers
            source          1..14
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 9; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
    |||||
Db 10 GCGGCGGCG 2

RESULT 13
BD068919/c
LOCUS      BD068919      14 bp      RNA      linear      PAT 27-AUG-2002
DEFINITION Enzymatic nucleic acid treatment of diseases or conditions related
           to levels of epidermal growth factor receptors.
ACCESSION  BD068919
VERSION     BD068919.1 GI:22614522
KEYWORDS    JP 2001511003-A/1759.
SOURCE      unidentified
ORGANISM    unclassified.
REFERENCE   1 (bases 1 to 14)
AUTHORS    Akhtar,S., Fell,P. and McSwiggen,J.A.
TITLE      Enzymatic nucleic acid treatment of diseases or conditions related
           to levels of epidermal growth factor receptors
JOURNAL    Patent: JP 2001511003-A 1759 07-AUG-2001;
           RIBOZYME PHARMACEUTICALS INC,ASTON UNIV
COMMENT     OS Unidentified
           PN JP 2001511003-A/1759
           PD 07-AUG-2001
           PF 14-JAN-1998 JP 1998532913
           PR 31-JAN-1997 US 60/036476,04-DEC-1997 US 08/985162 PI
           SAGHIR AKHTAR,PATRICIA FELL,JAMES A MCSWIGGEN PC
           C12N9/00,C07K14/71
           CC Strandedness: Single;
           CC Topology: Linear;
           CC Enzymatic nucleic acid treatment of diseases or conditions CC
           related to
           levels of epidermal growth factor receptors
           FH Key
           FT source
           FT 1..14
           FT Location/Qualifiers
           /organism="Unidentified".
           /organism="Unidentified"
           /mol_type="genomic RNA"
           /db_xref="taxon:32644"

ORIGIN
Query Match      100.0%; Score 9; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
    |||||
Db 10 GCGGCGGCG 2

RESULT 15
A88671
LOCUS      A88671      15 bp      DNA      linear      PAT 22-JAN-2000
DEFINITION Sequence 819 from Patent WO9833904.
ACCESSION  A88671
VERSION     A88671.1 GI:6737241
KEYWORDS

```



SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.  
 REFERENCE 1 (bases 1 to 15)  
 AUTHORS Brysch,W. and Schlingensiefen,K.  
 TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD  
 JOURNAL Patent: WO 9833904-A 819 06-AUG-1998;  
 BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)  
 FEATURES Location/Qualifiers  
 source 1..15  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 ORIGIN  
 Query Match 100.0%; Score 9; DB 6; Length 15;  
 Best Local Similarity 100.0%; Pred No. 3.8e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGGGCGG 9  
 |||||  
 Db 4 GCGGGGCGG 12

Search completed: October 28, 2004, 23:43:59  
 Job time : 50.3581 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: October 28, 2004, 20:36:09 ; Search time 93.8122 Seconds  
(without alignments)  
3495.894 Million cell updates/sec

Title: US-09-867-159a-6

Perfect score: 9

Sequence: 1 ggcggcgccg 9

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9	100.0	20	6	CF281215	14ETL--08
2	9	100.0	20	6	CF317946	HD--07-N0
3	9	100.0	20	6	CF339443	RCL1--04
4	9	100.0	22	2	BF979698	
5	9	100.0	23	8	AZ324338	1M0046B16
6	9	100.0	24	6	CF291636	14ROOT--0
7	9	100.0	26	6	CF292605	30DGS--01
8	9	100.0	27	6	CF339613	RCL1--05-
9	9	100.0	28	8	AZ503891	1M0343A24
10	9	100.0	29	6	CF327746	NACL--02-
11	9	100.0	30	2	BE539470	
12	9	100.0	31	2	BE904656	601498767
13	9	100.0	31	2	BF666846	602121169
14	9	100.0	31	4	BI156400	602919676
15	9	100.0	31	4	BI915569	603176924
16	9	100.0	31	9	CG720139	1119061A1
17	9	100.0	32	4	BI553076	
18	9	100.0	33	2	BE886705	603196045
19	9	100.0	33	6	CF299587	7LEAF--03
20	9	100.0	34	4	BI464279	603203394
21	9	100.0	34	9	AJ591955	Arabidops
22	9	100.0	34	9	CG728542	1119101E0
23	9	100.0	35	1	AL045094	DXF2p434L
24	9	100.0	35	4	EG122568	602353065

25 9 100.0 35 4 BG284367 602408473  
26 9 100.0 35 4 BG700656 602682274  
27 9 100.0 35 4 B1088801 602851610  
28 c 28 9 100.0 35 6 CF298077 7LEAF--01  
29 9 100.0 35 9 CG733046 1119153G0  
30 c 30 9 100.0 35 9 CL293918 0280349-0  
31 9 100.0 36 2 BF340538 602037175  
32 9 100.0 37 1 AA873697 0b36g11.s  
33 9 100.0 37 2 BF528080 603042827  
34 9 100.0 37 4 B1762138 603049308  
35 9 100.0 37 5 BQ605582 BRY\_1070  
36 c 36 9 100.0 37 8 A2758338 1M0550J17  
37 9 100.0 38 2 BF337026 602034416  
38 9 100.0 38 6 CF279323 14ETL--05  
39 9 100.0 39 2 BE871689 601449550  
40 c 40 9 100.0 39 7 D42385  
41 9 100.0 39 7 T49497 ya76807.x1  
42 c 42 9 100.0 39 8 AZ386752  
43 9 100.0 40 4 BG772038 1M0145G23  
44 c 44 9 100.0 40 6 CF280478 602721673  
45 9 100.0 40 9 CG729600 1119113E0

#### ALIGNMENTS

RESULT 1  
CF281215  
LOCUS  
DEFINITION  
14ETL--08-C23, 20 bp mRNA linear EST 14-AUG-2003  
Oryza sativa (japonica cultivar-group) cDNA clone 14ETL--08-C23,  
mRNA sequence.  
ACCESSION  
CF281215  
VERSION  
CF281215.1 GI:33658602  
KEYWORDS  
EST.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1 (bases 1 to 20)  
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,  
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.  
Large-scale Sequencing Analysis of Rice ESTs  
Unpublished (2003)  
Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyeonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.  
Location/Qualifiers  
1. .20  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:39947"  
/clone="14ETL--08-C23"  
/tissue\_type="leaf"  
/dev\_stage="14 days after germination"  
/lab\_host="E.coli DH10B"  
/clone\_lib="Rice etiolated leaf plasmid cDNA library  
(14ETL)"  
/note="vector: PCR4-TOP0; Site 1: EcoRI; mRNA was capped  
with oligoribonucleotides and then used as templates for  
RT-PCR."

FEATURES  
source

#### ORIGIN

Query Match 100.0%; Score 9; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GCGGCGGCG 9
      |||||
Db      1 GCGGCGGCG 9

RESULT 2
CF317946
LOCUS   HD-07-N06.g1 OsHDA1-overexpressing transgenic rice plasmid cDNA
DEFINITION
library (HD) Oryza sativa (japonica cultivar-group) cDNA clone
HD-07-N06, mRNA sequence.
ACCESSION
CF317946
VERSION
CF317946.1 GI:33689707
KEYWORDS
EST.
SOURCE  Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnam@gbio.com, bhnam@bio.myongji.ac.kr.
Location/Qualifiers
1..20
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="RCL1--04-003"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 30 days"
/lab_host="E.coli SOLR"
/clone_lib="Regenerated callus lambda phage cDNA library
(RCL1)"
/notes="Vector: pBluescript SK(+); Site 1: SstI; Site 2:
XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5'
end with SstI and 3' end with XhoI site. Callus was
induced on 2N6 media for 30 days and cultured for 36hrs on
regenerated media"

FEATURES
source
Query Match 100.0%; Score 9; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGCGGCG 9
      |||||
Db      1 GCGGCGGCG 18

RESULT 4
BF979698
LOCUS   602288551F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4374157 5',
DEFINITION
mRNA sequence.
ACCESSION
BF979698
VERSION
BF979698.2 GI:12388195
KEYWORDS
EST.
SOURCE   Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 22)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   On Jan 17, 2001 this sequence version replaced gi:12346913.
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLNL0037 row: f column: 14
High quality sequence stop: 22.
Location/Qualifiers
1..22
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source
Query Match 100.0%; Score 9; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGCGGCG 9
      |||||
Db      1 GCGGCGGCG 11

RESULT 3
CF339443
LOCUS   RCL1--04-003.g1 Regenerated callus lambda phage cDNA library (RCL1)
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone RCL1--04-003,
mRNA sequence.
ACCESSION
CF339443
VERSION
CF339443.1 GI:33927271
KEYWORDS
EST.
SOURCE  Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

```

```

/db_xref="taxon:9606"
/clone="IMAGE:4374157"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtccag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to R0T 5. This is a
primary library enriched for full-length clones and
constructed using the cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match          100.0%; Score 9; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGGCG 9
    |||||
Db 6 GCGGCGGGCG 14

RESULT 5
AZ324328
LOCUS
DEFINITION
IM0046B:6F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0046B16 F, genomic survey sequence.
ACCESSION
AZ324328
VERSION
AZ324328.1 GI:10379937
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 23)
Dunn,D., Aoyagi,A., Barber,M., Bescorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,R., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5506
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0046 row: B column: 16
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 23.

FEATURES
source
1..23
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0046B16"
/sex="Male"
/lab_host="E. Coli strain Xli0-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnarses/). The DNA
was hydrodynamically sheared by repeated passage through a

```

```

0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (G14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match          100.0%; Score 9; DB 8; Length 23;
Best Local Similarity 100.0%; Pred. No. 6.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGGCG 9
    |||||
Db 8 GCGGCGGGCG 16

RESULT 6
CF291636
LOCUS
DEFINITION
14ROOT--02-C09.g1 Rice root plasmid cDNA library (14ROOT) Oryza
sativa (japonica cultivar-group) cDNA clone 14ROOT--02-C09, mRNA
sequence.
ACCESSION
CF291636
VERSION
CF291636.1 GI:33660669
KEYWORDS
EST.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriartoideae; Oryzaceae; Oryza.
1 (bases 1 to 24)
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
Contact: Nam B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Yeongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
source
1..24
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="14ROOT--02-C09"
/tissue_type="root"
/dev_stage="14 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice root plasmid cDNA library (14ROOT)"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match          100.0%; Score 9; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.2e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGGCG 9
    |||||

```

```

Db      1 GCGGCGGCG 9

RESULT 7
CF292605/c
LOCUS      26 bp      mRNA      linear      EST 14-AUG-2003
DEFINITION 30DGS--01-H22.g1 Rice leaf plasmid cDNA library I (30DGS) Oryza
           sativa (japonica cultivar-group) cDNA clone 30DGS--01-H22, mRNA
           sequence.
ACCESSION  CF292605
VERSION     CF292605.1 GI:33661638
KEYWORDS   EST.
SOURCE     Oryza sativa (japonica cultivar-group)
           Oryza sativa (japonica cultivar-group)
           Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 26)
           Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
           Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
           Large-scale Sequencing Analysis of Rice ESTs
           Unpublished (2003)
JOURNAL
COMMENT    Contact: Nahm B.H.
           Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
           of Bioscience and Bioinformatics, Myongji University
           Yongin, Gyeonggi, Korea
           Tel: 82 31 330 6193
           Fax: 82 31 321 6355
           Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.
           Location/Qualifiers
           1..26
           /organism="Oryza sativa (japonica cultivar-group)"
           /mol_type="mRNA"
           /cultivar="Nackdong"
           /db_xref="taxon:39947"
           /clone="RCL1--05-G12"
           /tissue_type="callus"
           /dev_stage="proliferated callus on 2N6 media for 30 days"
           /lab_host="E.coli SOLR"
           /clone_lib="Regenerated callus lambda phage cDNA library
           (RCL1)"
           /note="vector: pBluescript SK(+); Site 1: SstI; Site 2:
           XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5'
           end with SstI and 3' end with XhoI site. Callus was
           induced on 2N6 media for 30 days and cultured for 36hrs on
           regenerated media"

ORIGIN
Query Match      100.0%; Score 9; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCGGCGGCG 9
        |||||
Db      26 GCGGCGGCG 18

RESULT 9
AZ503891
LOCUS      28 bp      DNA      linear      GSS 05-OCT-2000
DEFINITION 1M0343A24R Mouse 10kb plasmid UGCLM library Mus musculus genomic
           clone UGCLM0343A24 R, genomic survey sequence.
ACCESSION  AZ503891
VERSION     AZ503891.1 GI:10685207
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
           1 (bases 1 to 28)
REFERENCE  1
AUTHORS   Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
           Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
           Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
           Niederhausern, A. and Wright, D., Weiss, R.
           Mouse whole genome scaffolding with paired end reads from 10kb
           plasmid inserts
           Unpublished (2000)
JOURNAL
COMMENT    Contact: Robert B. Weiss
           University of Utah Genome Center
           University of Utah
           Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
           84112, USA
           Tel: 801 585 5606
           Fax: 801 585 7177
           Email: ddunn@genetics.utah.edu
           Insert Length: 10000 Std Error: 0.00
           Plate: 0343 row: A column: 24
           Seq primer: CACACAGAACACGCTATGACC
           Class: plasmid ends
           High quality sequence stop: 28.
           Location/Qualifiers
           1..28
           /organism="Mus musculus"
           /mol_type="Genomic DNA"
           /strain="C57BL/6J"

FEATURES
source
Query Match      100.0%; Score 9; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 6.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCGGCGGCG 9
        |||||
Db      15 GCGGCGGCG 7

RESULT 8
CF339613/c
LOCUS      27 bp      mRNA      linear      EST 18-AUG-2003
DEFINITION RCL1--05-G12.g1 Regenerated callus lambda phage cDNA library (RCL1)
           Oryza sativa (japonica cultivar-group) cDNA clone RCL1--05-G12,
           mRNA sequence.
ACCESSION  CF339613
VERSION     CF339613.1 GI:33827599
KEYWORDS   EST.
SOURCE     Oryza sativa (japonica cultivar-group)
           Oryza sativa (japonica cultivar-group)
           Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 27)
           Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
           Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
           Large-scale Sequencing Analysis of Rice ESTs
           Unpublished (2003)
JOURNAL
COMMENT

```

```

/db_xref="taxon:10090"
/clone="UUGC1M0343A24"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: pW42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pW42 (g[4732114]gb[AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 9; DB 8; Length 28;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9
Db 19 GCGGCGGCG 27

RESULT 10
CF327746
LOCUS
DEFINITION
NACL--02-F18.g1 Rice callus plasmid cDNA library (NACL) Oryza sativa (japonica cultivar-group) cDNA clone NACL--02-F18, mRNA sequence.
25 bp mRNA linear EST 18-AUG-2003

ACCESSION
CF327746
VERSION
CF327746.1 GI:33803742
KEYWORDS
EST.
SOURCE
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 29)
AUTHORS
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
CONTACT: Nahm B.H.
Genomics and Genetics Institute, Greengene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.

FEATURES
Location/Qualifiers
1..29
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="NACL--02-F18"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 30 days"
/lab_host="E.coli DH10B"
/clone_lib="Rice callus plasmid cDNA library (NACL)"
/notes="Vector: PCR4-TOPO; Site_1: EcoRI; mRNA was capped
```

with oligoribonucleotides and then used as templates for RT-PCR."

```

ORIGIN
Query Match 100.0%; Score 9; DB 6; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.9e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9
Db 2 GCGGCGGCG 10

RESULT 11
BE539470
LOCUS
DEFINITION
601060134F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3446720 5', mRNA sequence.
30 bp mRNA linear EST 09-AUG-2000

ACCESSION
BE539470
VERSION
BE539470.1 GI:9768115
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 30)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8419 row: c column: 09
High quality sequence stop: 30.
Location/Qualifiers
1..30
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3446720"
/cell_line="MGC36"
/lab_host="DH10B"
/clone_lib="NIH_MGC_10"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies."
```

## ORIGIN

Query Match 100.0%; Score 9; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 5.9e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
Db 14 GCGGCGGCG 22

## RESULT 12

BE904656  
LOCUS  
DEFINITION  
601498767F1 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE:3900665 5', mRNA sequence.  
30 bp mRNA linear EST 20-OCT-2000  
ACCESSION  
BE904656  
VERSION  
BE904656.1 GI:10397135  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 30)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9700 row: e column: 18
High quality sequence stop: 30.
FEATURES
source
1..30
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 70"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 100.0%; Score 9; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.9e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGCGCG 9
Db 3 GCGCGCGCG 11
RESULT 13
BF666846 602121169F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278257 5',
DEFINITION mRNA sequence.
ACCESSION BF666846
VERSION BF666846
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 31)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM101 row: j column: 18
High quality sequence stop: 31.
FEATURES
source
1..31
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone_lib="NIH MGC 56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcccttggcc); Site 2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
ORIGIN
Query Match 100.0%; Score 9; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 5.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGCGCG 9
Db 18 GCGCGCGCG 26
RESULT 14
BI156400 60291967F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5060011 5',
LOCUS mRNA sequence.
DEFINITION BI156400
ACCESSION BI156400
VERSION BI156400.1 GI:14616401
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 31)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1162 row: g column: 20.
FEATURES
Location/Qualifiers
1..31
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:5060011"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
ORIGIN
Query Match 100.0%; Score 9; DB 4; Length 31;

```



Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
Db 21 GCGGCGGCG 29

RESULT 15

BI915569  
LOCUS 603176924F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5241437 5',  
DEFINITION mRNA sequence.

ACCESSION BI915569  
VERSION BI915569.1 GI:16179572  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 31)  
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: L1AM1608 row: 0 column: 06  
High quality sequence stop: 31.

FEATURES

source  
1..31  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5241437"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_121"  
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3  
fetal brains, female age 20 weeks, female age 24 weeks,  
and male age 26 weeks. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 Kb, insert size range  
0.7-3.5 Kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 017. Note:  
this is a NIH\_MGC Library."

ORIGIN

Query Match 100.0%; Score 9; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
Db 2 GCGGCGGCG 10

Search completed: October 29, 2004, 01:43:39  
Job time : 101.812 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: October 28, 2004, 19:39:58 ; Search time 11.1616 Seconds  
(without alignments)  
4232.808 Million cell updates/sec

Title: US-09-867-159A-6  
Perfect score: 9  
Sequence: 1 gggcgggcg 9  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4134886 seqs, 2624710521 residues  
Total number of hits satisfying chosen parameters: 8269772  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04: \*  
1: Geneseq1980s: \*  
2: Geneseq1990s: \*  
3: Geneseq2000s: \*  
4: Geneseq2001as: \*  
5: Geneseq2001bs: \*  
6: Geneseq2002as: \*  
7: Geneseq2002bs: \*  
8: Geneseq2003as: \*  
9: Geneseq2003bs: \*  
10: Geneseq2003cs: \*  
11: Geneseq2003ds: \*  
12: Geneseq2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	%		Length	DB	ID	Description		
		Match	Query						
1	9	100.0	10	3	AAZ86062		Aaz86062 Metastat		
2	9	100.0	10	3	AAZ80968		Aaz80968 Metastat		
3	9	100.0	10	8	ACC78769		Acc78769 Normal e		
4	9	100.0	11	6	ABV68291		Abv68291 Human sh		
5	9	100.0	11	10	ADJ94151		Adj94151 Oligonu		
6	9	100.0	11	12	ADL16098		Adl16098 Neisser		
C 7	9	100.0	11	12	ADQ33885		Adq33885 Human f		
C 8	9	100.0	11	12	ADQ33920		Adq33920 Human f		
C 9	9	100.0	12	2	AAV05921		Aav05921 Transla		
10	9	100.0	12	5	ABI20277		Abi20277 Oligonu		
11	9	100.0	12	5	ABI20278		Abi20278 Oligonu		
12	9	100.0	12	6	AAS13692		Aas13692 DNA enc		
13	9	100.0	12	6	AAL41283		Aal41283 Oligonu		
14	9	100.0	12	6	ABQ08034		Abq08034 Cysteine		
15	9	100.0	12	8	ABX80021		Abx80021 EST poly		
16	9	100.0	12	10	ADE14346		Ade14346 Optiner		
C 17	9	100.0	12	12	ADH09454		Adh09454 Propion		
C 18	9	100.0	14	2	AAV98979		Aav98979 Human E		
C 19	9	100.0	14	2	AAV98980		Aav98980 Human E		
C 20	9	100.0	14	3	AAA26132		Aaa26132 Oestrog		
C 21	9	100.0	14	3	AAA26133		Aaa26133 Oestrog		

22	9	100.0	15	2	AAT86420	9	100.0	15	2	AAT86420	9	100.0	15	2	AAT86420	9	100.0	15	2	AAT86420
23	9	100.0	15	2	AAV49230	9	100.0	15	2	AAV49230	9	100.0	15	2	AAV49230	9	100.0	15	2	AAV49230
24	9	100.0	15	3	AAA40336	9	100.0	15	3	AAA40336	9	100.0	15	3	AAA40336	9	100.0	15	3	AAA40336
25	9	100.0	15	4	AAF45313	9	100.0	15	4	AAF45313	9	100.0	15	4	AAF45313	9	100.0	15	4	AAF45313
26	9	100.0	15	4	AAF45317	9	100.0	15	4	AAF45317	9	100.0	15	4	AAF45317	9	100.0	15	4	AAF45317
27	9	100.0	15	4	AAF45371	9	100.0	15	4	AAF45371	9	100.0	15	4	AAF45371	9	100.0	15	4	AAF45371
28	9	100.0	15	4	AAF45310	9	100.0	15	4	AAF45310	9	100.0	15	4	AAF45310	9	100.0	15	4	AAF45310
29	9	100.0	15	4	AAF45375	9	100.0	15	4	AAF45375	9	100.0	15	4	AAF45375	9	100.0	15	4	AAF45375
30	9	100.0	15	4	AAF45318	9	100.0	15	4	AAF45318	9	100.0	15	4	AAF45318	9	100.0	15	4	AAF45318
31	9	100.0	15	4	AAF45373	9	100.0	15	4	AAF45373	9	100.0	15	4	AAF45373	9	100.0	15	4	AAF45373
32	9	100.0	15	4	AAF45312	9	100.0	15	4	AAF45312	9	100.0	15	4	AAF45312	9	100.0	15	4	AAF45312
33	9	100.0	15	4	AAF45315	9	100.0	15	4	AAF45315	9	100.0	15	4	AAF45315	9	100.0	15	4	AAF45315
34	9	100.0	15	4	AAF45316	9	100.0	15	4	AAF45316	9	100.0	15	4	AAF45316	9	100.0	15	4	AAF45316
35	9	100.0	15	4	AAF45367	9	100.0	15	4	AAF45367	9	100.0	15	4	AAF45367	9	100.0	15	4	AAF45367
36	9	100.0	15	4	AAF45376	9	100.0	15	4	AAF45376	9	100.0	15	4	AAF45376	9	100.0	15	4	AAF45376
37	9	100.0	15	4	AAF45314	9	100.0	15	4	AAF45314	9	100.0	15	4	AAF45314	9	100.0	15	4	AAF45314
38	9	100.0	15	4	AAF45311	9	100.0	15	4	AAF45311	9	100.0	15	4	AAF45311	9	100.0	15	4	AAF45311
39	9	100.0	15	4	AAF45368	9	100.0	15	4	AAF45368	9	100.0	15	4	AAF45368	9	100.0	15	4	AAF45368
40	9	100.0	15	4	AAF45369	9	100.0	15	4	AAF45369	9	100.0	15	4	AAF45369	9	100.0	15	4	AAF45369
41	9	100.0	15	4	AAF45370	9	100.0	15	4	AAF45370	9	100.0	15	4	AAF45370	9	100.0	15	4	AAF45370
42	9	100.0	15	4	AAF45372	9	100.0	15	4	AAF45372	9	100.0	15	4	AAF45372	9	100.0	15	4	AAF45372
43	9	100.0	15	4	AAF45374	9	100.0	15	4	AAF45374	9	100.0	15	4	AAF45374	9	100.0	15	4	AAF45374
44	9	100.0	15	4	AAF45319	9	100.0	15	4	AAF45319	9	100.0	15	4	AAF45319	9	100.0	15	4	AAF45319
45	9	100.0	15	6	AAS19754	9	100.0	15	6	AAS19754	9	100.0	15	6	AAS19754	9	100.0	15	6	AAS19754

ALIGNMENTS

RESULT 1  
AAZ86062  
ID AAZ86062 standard; DNA; 10 BP.  
XX  
AC AAZ86062;  
XX  
DT 07-APR-2000 (first entry)  
XX  
DE Metastatic breast tumour cell downregulated transcript tag #5296.  
XX  
KW Human; metastatic breast tumour tissue; breast cancer; tag; primer;  
non-metastatic breast tumour tissue; gene therapy; anticancer;  
antimetastatic; vaccine; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9965928-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 18-JUN-1999; 99WO-US013647.  
XX  
PR 19-JUN-1998; 98US-0089853P.  
PR 19-JUN-1998; 98US-0089977P.  
PR 19-JUN-1998; 98US-0090039P.  
PR 19-JUN-1998; 98US-0090040P.  
XX  
XX (GENZ ) GENZYME CORP.  
PA (ROBE/) ROBERTS B L.  
PA (SHAN/) SHANKARA S.  
XX  
PI Roberts BL, Shankara S;  
XX  
XX WPI; 2000-106079/09.  
XX  
XX Isolated polynucleotides differentially expressed between metastatic and  
non-metastatic breast cancer cells, useful for diagnosis, prevention and  
treatment of cancer.  
XX  
XX Claim 1; page 199; 219pp; English.  
XX  
XX AAZ80767 to AAZ83941 represent tags corresponding to distinct transcripts  
that are preferentially transcribed in the metastatic breast tumour

CC tissue (i.e. are upregulated in metastatic breast tumour cells). AAZ83942  
 CC to AAZ86677 represent tags corresponding to distinct transcripts that are  
 CC preferentially transcribed in the primary or non-metastatic breast tumour  
 CC tissue (i.e. are downregulated in metastatic breast tumour cells). These  
 CC transcripts can be used for diagnosis, prognosis, monitoring and  
 CC treatment of breast cancer, particularly where metastatic. Diagnosis is  
 CC by standard immunoassays or hybridisation/amplification reactions.  
 CC Compounds that modulate expression of the transcripts are potentially  
 CC useful for treatment of (metastatic) breast cancer, while promoters from  
 CC the transcripts are used to direct expression, in selected cell types, of  
 CC e.g. therapeutic genes (also ribozymes or antisense sequences),  
 CC particularly an antigen-encoding sequence for use in gene or cell-based  
 CC vaccines. Polypeptides encoded by the transcripts are also useful in  
 CC antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic  
 CC agents. Host cells that produce the polypeptides can be used to expand  
 CC and isolate populations of educated, antigen-specific immune effector  
 CC cells, e.g. cytotoxic T lymphocytes, and these used for adoptive  
 CC immunotherapy  
 XX  
 SQ Sequence 10 BP; 0 A; 3 C; 7 G; 0 T; 0 U; 0 Other;

Query Match 100.0%; Score 9; DB 3; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 |||||  
 Db 1 GCGGCGGCG 9

## RESULT 2

AAZ80968  
 ID AAZ80968 standard; DNA; 10 BP.  
 AC AAZ80968;  
 XX  
 DT 07-APR-2000 (first entry)  
 XX  
 DE Metastatic breast tumour cell upregulated transcript tag #202.  
 XX  
 KW Human; metastatic breast tumour tissue; breast cancer; tag; primer;  
 KW non-metastatic breast tumour tissue; gene therapy; anticancer;  
 KW antimetastatic; vaccine; diagnosis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 DN W0995928-A2.  
 XX  
 PD 23-DEC-1999.  
 XX  
 PF 18-JUN-1999; 99WO-US013647.  
 XX  
 PR 19-JUN-1998; 98US-0089853P.  
 PR 19-JUN-1998; 98US-0089997P.  
 PR 19-JUN-1998; 98US-0090039P.  
 PR 19-JUN-1998; 98US-0090040P.  
 PR 19-JUN-1998; 98US-0090041P.  
 XX  
 PA (GENZ ) GENZYME CORP.  
 PA (ROBE/) ROBERTS B.L.  
 PA (SHAN/) SHANKARA S.  
 XX  
 PI Roberts BL, Shankara S;  
 XX  
 DR WPI; 2000-106079/09.  
 XX  
 XX Isolated polynucleotides differentially expressed between metastatic and  
 PT non-metastatic breast cancer cells, useful for diagnosis, prevention and  
 PT treatment of cancer.  
 XX  
 PS Claim 1; Page 63; 219pp; English.  
 XX

CC AAZ80767 to AAZ83941 represent tags corresponding to distinct transcripts  
 CC that are preferentially transcribed in the metastatic breast tumour  
 CC tissue (i.e. are upregulated in metastatic breast tumour cells). AAZ83942  
 CC to AAZ86677 represent tags corresponding to distinct transcripts that are  
 CC preferentially transcribed in the primary or non-metastatic breast tumour  
 CC tissue (i.e. are downregulated in metastatic breast tumour cells). These  
 CC transcripts can be used for diagnosis, prognosis, monitoring and  
 CC treatment of breast cancer, particularly where metastatic. Diagnosis is  
 CC by standard immunoassays or hybridisation/amplification reactions.  
 CC Compounds that modulate expression of the transcripts are potentially  
 CC useful for treatment of (metastatic) breast cancer, while promoters from  
 CC the transcripts are used to direct expression, in selected cell types, of  
 CC e.g. therapeutic genes (also ribozymes or antisense sequences),  
 CC particularly an antigen-encoding sequence for use in gene or cell-based  
 CC vaccines. Polypeptides encoded by the transcripts are also useful in  
 CC antibodies (Ab). Ab are used to detect the polypeptides or as therapeutic  
 CC agents. Host cells that produce the polypeptides can be used to expand  
 CC and isolate populations of educated, antigen-specific immune effector  
 CC cells, e.g. cytotoxic T lymphocytes, and these used for adoptive  
 CC immunotherapy  
 XX

SQ Sequence 10 BP; 1 A; 3 C; 6 G; 0 T; 0 U; 0 Other;

Query Match 100.0%; Score 9; DB 3; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 |||||  
 Db 1 GCGGCGGCG 9

## RESULT 3

ACC78769  
 ID ACC78769 standard; DNA; 10 BP.  
 AC ACC78769;  
 XX  
 DT 02-SEP-2003 (first entry)  
 XX  
 DE Normal estrogen responsive cells.derived SAGE tag.  
 XX  
 KW ERE; reporter construct; estrogen response element; cytostatic; rat;  
 KW gene therapy; breast cancer; SAGE; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W02003042364-A2.  
 XX  
 PD 22-MAY-2003.  
 XX  
 PF 08-NOV-2002; 2002WO-US035901.  
 XX  
 PR 09-NOV-2001; 2001US-0338136P.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Polyak K, Pankaj S;  
 XX  
 DR WPI; 2003-449570/42.  
 XX  
 XX New reporter construct for identifying and isolating estrogen-responsive  
 PT cells comprises an estrogen response segment, a promoter segment and a  
 PT nucleotide sequence that encodes a reporter polypeptide.  
 XX  
 PS Example 4; Page 32; Sipp; English.  
 XX  
 CC The invention relates to a reporter construct comprising: (a) an estrogen  
 CC response segment having 5 or more estrogen response elements (ERE); (b) a  
 CC promoter segment having at least one promoter nucleic acid sequence; and  
 CC (c) a nucleotide sequence that encodes a reporter polypeptide, where the  
 CC nucleotide sequence is operably linked to the promoter segment and the

CC estrogen response segment. The reporter construct and vector are useful  
CC in identifying and isolating estrogen-responsive cells. The methods are  
CC useful in inhibiting the proliferation or survival of estrogen-responsive  
CC breast cancer cells or in enhancing the proliferation or survival of  
CC estrogen-receptor non-expressing, estrogen-non-responsive cells.  
CC Sequences ACC78740-75 represent SAGE tags for transcripts specifically or  
CC most abundantly expressed in normal estrogen responsive cells  
XX  
SQ Sequence 10 BP; 0 A; 3 C; 7 G; 0 T; 0 U; 0 Other;

Query Match 100.0%; Score 9; DB 8; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
Db 1 GCGGCGGCG 9

RESULT 4  
ABV68291  
ID ABV68291 standard; cDNA; 11 BP.

XX AC ABV68291;  
XX  
DT 21-OCT-2002 (first entry)

XX DE Human skin EST 6077.  
XX Human; skin; dermatological; vulnary; antipsoriatic; antiseborrhaic;  
KW immunosuppressive; antiinflammatory; cytostatic; SAGE; neurodermatitis;  
KW psoriasis; dermatitis; skin cancer; EST; expressed sequence tag; ss.

XX OS Homo sapiens.  
XX  
XX WO200253774-A2.  
XX  
PD 11-JUL-2002.

XX PF 20-DEC-2001; 2001WO-EP015179.  
XX  
XX 03-JAN-2001; 2001DE-01000127.  
XX  
XX (HENK ) HENKEL KGAA.

XX PI Petersohn D, Conradt M, Hofmann K;  
XX  
XX WPI; 2002-590638/63.

XX  
PT In vitro identification of skin-expressed genes, useful for determining  
PT homeostasis and identifying cosmetic or pharmaceutical agents against  
PT e.g. skin cancer.

XX PS Disclosure; Page 193; 1345pp; German.

XX CC The invention relates to in vitro identification (M1) of genes expressed  
CC in the skin of humans or animals by subjecting a mixture of genetically  
CC encoded factors from skin, to serial analysis of gene expression (SAGE)  
CC so as to identify skin-expressed genes and quantify their expression.  
CC (M1) is useful for identifying genes involved in skin homeostasis; to  
CC determine skin homeostasis and to test agent (A) that maintains or  
CC promotes skin homeostasis or that can be used for treating skin  
CC disorders, specifically neurodermatitis; sunburn; psoriasis; scleroderma;  
CC ichthyosis; atopic dermatitis; acne; seborrhea; lupus erythematosus;  
CC rosacea; melanoma; basal cell carcinoma; and carcinoma or sarcoma of the  
CC skin. The present sequence is that of a human expressed sequence tag  
CC (EST) of the invention

XX SQ Sequence 11 BP; 1 A; 3 C; 7 G; 0 T; 0 U; 0 Other;  
Query Match 100.0%; Score 9; DB 6; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
Db 1 GCGGCGGCG 9

RESULT 5  
ADJ94151  
ID ADJ94151 standard; DNA; 11 BP.

XX AC ADJ94151;  
XX  
DT 06-MAY-2004 (first entry)

XX DE Oligonucleotide of the invention #18.  
XX ds; linear polynucleotide; cation; telomere repeat.  
XX Unidentified.

XX JN JP2003169676-A.  
XX  
PD 17-JUN-2003.

XX PF 05-DEC-2001; 2001JP-00371975.  
XX  
XX 05-DEC-2001; 2001JP-00371975.

XX PA (TAKE/) TAKENAKA S.  
XX (TUMK-) TUM KENYUSHO KK.  
XX  
XX WPI; 2003-639772/61.

XX PT Quantitative method of cation for clinical laboratory test, involves  
PT adding linear polynucleotide containing telomere repeat sequence with  
PT donor and acceptor at edge portions, to solution containing cation.  
XX  
XX Disclosure; Page 21; 8pp; Japanese.

XX CC The invention relates to a novel quantitative method for adding linear  
CC polynucleotide to solution containing a cation. The linear polynucleotide  
CC contains telomere repeat sequence with donor at one edge portion and  
CC acceptor at another edge portion. The optical detection is performed  
CC using this linear polynucleotide. The method of the invention is useful  
CC in medical fields such as clinical laboratory test and fields with cation  
CC fixed assay. The present sequence is used in the exemplification of the  
CC invention.

XX SQ Sequence 11 BP; 0 A; 3 C; 8 G; 0 T; 0 U; 0 Other;

Query Match 100.0%; Score 9; DB 10; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
Db 2 GCGGCGGCG 10

RESULT 6  
ADL16098/c  
ID ADL16098 standard; DNA; 11 BP.

XX AC ADL16098;  
XX  
DT 06-MAY-2004 (first entry)

XX DE Neisseria meningitidis lgtG "fixed" mutant gene disrupted polyc tract.  
XX  
XX Lipooligosaccharide immunotype; LOS immunotype; serogroup B;  
KW phase variation; fixed immunotype; homopolymetric nucleotide tract;  
KW vaccine; immunostimulant; meningococcal disease; Neisserial disease;  
KW mutant; lgtG; disrupted polyc tract; fixed; constitutive expression; ds.

```

XX OS Neisseria meningitidis; strain 35E.
XX OS Synthetic.
XX FH Key
XX FT mutation
XX FT Location/Qualifiers
XX FT replace(2,C)
XX FT /*tag= a
XX FT /note= "This base is C in the wild-type lgtG gene"
XX FT mutation
XX FT replace(S,C)
XX FT /*tag= b
XX FT /note= "This base is C in the wild-type lgtG gene"
XX FT mutation
XX FT replace(8,C)
XX FT /*tag= c
XX FT /note= "This base is C in the wild-type lgtG gene"
XX FT
XX WO2004015099-A2.
XX PN
XX PD
XX PD 19-FEB-2004.
XX XX
XX PF 31-JUL-2003; 2003WO-EP008569.
XX XX
XX PR 02-AUG-2002; 2002GB-00018035.
XX PR 02-AUG-2002; 2002GB-00018036.
XX PR 02-AUG-2002; 2002GB-00018037.
XX PR 02-AUG-2002; 2002GB-00018051.
XX PR 30-AUG-2002; 2002GB-00020197.
XX PR 30-AUG-2002; 2002GB-00020199.
XX PR 01-NOV-2002; 2002GB-00025524.
XX PR 01-NOV-2002; 2002GB-00025531.
XX PR 24-DEC-2002; 2002GB-00030164.
XX PR 24-DEC-2002; 2002GB-00030168.
XX PR 24-DEC-2002; 2002GB-00030170.
XX PR 05-MAR-2003; 2003GB-00005028.
XX XX
XX PA (GLAX ) GLAXOSMITHKLINE BIOLOGICALS SA.
XX PA (UYQU ) UNIV QUEENSLAND.
XX XX
XX PI Biemans R, Denoel P, Feron C, Goraj K, Jennings MP, Poolman J;
XX PI Weynants V;
XX XX
XX DR WPI; 2004-180668/17.
XX XX
XX PS Example 3; Page 28; 42pp; English.
XX CC
XX CC The invention relates to a process for making a genetically engineered
XX CC Neisserial strain (preferably Neisseria meningitidis serogroup B) in
XX CC which the lipooligosaccharide (LOS) immunotype is fixed or locked. A
XX CC feature of the meningococcal LOS is the reversible, high frequency
XX CC switching of expression (phase variation) of terminal LOS structures,
XX CC which is an obstacle to the development of a cross-protective vaccine
XX CC based on the use of LOS as the antigen. The process of the invention
XX CC involves engineering a Neisserial strain such that the homopolymERIC
XX CC nucleotide tract of a phase variable LOS synthesis gene (specifically
XX CC lgtA or lgtG) is reduced in length ( whilst maintaining the open reading
XX CC frame), resulting in gene expression which is less phase variable. The
XX CC method of the invention can be used to produce a Neisserial strain with a
XX CC fixed L2 or L3 immunotype, which can be used in the manufacture of
XX CC vaccines (particularly multivalent vaccines) against neisserial disease,
XX CC especially meningococcal disease. The present sequence represents the
XX CC disrupted polyc tract of the constitutively expressed Neisseria
XX CC meningitidis lgtG "fixed" mutant gene (ADL16103).
XX XX
XX SQ Sequence 11 BP; 0 A; 8 C; 3 G; 0 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 9; DB 12; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 1.6e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GCGGCGGCG 9
XX Db 9 GCGGCGGCG 1
XX
XX RESULT 8
XX ID ADQ33920 standard; DNA; 11 BP.
XX
XX RESULT 7
XX ID ADQ33885/c
XX AC ADQ33885;
XX DT 23-SEP-2004 (first entry)
XX DE Human facial skin-associated DNA fragment SEQ ID NO 1975.
XX KW facial skin; human; serial analysis of gene expression; SAGE;
XX KW homeostasis; biochip; cosmetic; pharmaceutical; ds.
XX OS Homo sapiens.
XX XX
XX PN DE10260928-A1.
XX PD 08-JUL-2004.
XX XX
XX PF 20-DEC-2002; 2002DE-01060928.
XX PR 20-DEC-2002; 2002DE-01060928.
XX XX (HENK ) HENKEL KGAA.
XX PI Petersohn D, Schlotmann K, Gassenmeier T, Holtkoetter O;
XX PI Conradt M, Hofmann K;
XX XX
XX DR WPI; 2004-518855/50.
XX XX
XX PT In vitro identification of genes important for facial skin, useful for
XX PT assessing homeostasis and in screening for pharmaceutical or cosmetic
XX PT agents, based on differential expression analysis.
XX PS Claim 5; SEQ ID NO 1975; 577pp; German.
XX CC
XX CC This invention describes a novel in vitro method for identifying genes
XX CC that are significant for facial skin in humans. The method comprises
XX CC recovering, from facial skin, a first mixture of genetically expressed
XX CC (transcribed and optionally translated) factors (i.e. proteins, mRNA or
XX CC their fragments), recovering a second, similar mixture from some other
XX CC human tissue, preferably skin from a protected area, especially from the
XX CC breast and subjecting the mixtures to serial analysis of gene expression
XX CC (SAGE) to identify those genes for which expression is markedly different
XX CC between facial skin and the other tissue. The invention also describes an
XX CC in vitro method for determining homeostasis of human facial skin; a test
XX CC kit which comprises a solid support (flexible or rigid) on which are
XX CC immobilised probes that bind specifically to the factors of interest and
XX CC a biochip for determining homeostasis of human facial skin. The products
XX CC of the invention are also used in a method which determines activity of
XX CC cosmetic and pharmaceutical agents for use against disorders or
XX CC disturbances of the homeostasis of human skin and a screening method for
XX CC identifying cosmetic and pharmaceutical agents. The method allows
XX CC identification of as many as possible of the genes important for facial
XX CC skin and thus of a very wide range of potential therapeutic and cosmetic
XX CC agents. ADQ31911-ADQ35111 represent human DNA Tag fragments used to
XX CC identify the facial skin-associated genes described in the invention.
XX SQ Sequence 11 BP; 0 A; 7 C; 4 G; 0 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 9; DB 12; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 1.6e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GCGGCGGCG 9
XX Db 11 GCGGCGGCG 3
XX
XX RESULT 8
XX ID ADQ33920 standard; DNA; 11 BP.
XX

```



```

XX PF 06-APR-2001; 2001WO-IB000713.
XX PS
XX PR 07-APR-2000; 2000DE-01019173.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2001-657177/75.
XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single-nucleotide polymorphisms and cytosine
PT methylation status.
XX PS Claim 1; SEQ ID NO 320250; 29pp + Sequence Listing; German.
XX CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation. ABC00010
CC -ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and ABI00010-ABI82073
CC represent the oligomers described in the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 12 BP; 0 A; 3 C; 8 G; 1 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 9; DB 5; Length 12;
XX Best Local Similarity 100.0%; Pred. NO. 1.6e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCGGCGGCG 9
XX Db 3 GCGGCGGCG 11
XX
XX RESULT 11
XX ABI20278
XX ID ABI20278 standard; DNA; 12 BP.
XX AC ABI20278;
XX XX 22-FEB-2002 (first entry)
XX DE Oligonucleotide primer SEQ ID NO 320251 for detecting SNP TSC0029618.
XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX OS Homo sapiens.
XX XX WO200177384-A2.
XX PN 18-OCT-2001.
XX PD
XX PF 06-APR-2001; 2001WO-IB000713.
XX PR 07-APR-2000; 2000DE-01019173.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2001-657177/75.
XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single-nucleotide polymorphisms and cytosine
PT methylation status.
XX PS Claim 1; SEQ ID NO 320251; 29pp + Sequence Listing; German.
XX CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation. ABC00010
CC -ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and ABI00010-ABI82073
CC represent the oligomers described in the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 12 BP; 0 A; 3 C; 8 G; 1 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 9; DB 5; Length 12;
XX Best Local Similarity 100.0%; Pred. NO. 1.6e+05;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GCGGCGGCG 9
XX Db 3 GCGGCGGCG 11
XX
XX RESULT 12
XX AAS13692/c
XX ID AAS13692 standard; DNA; 12 BP.
XX AC AAS13692;
XX XX 29-JAN-2002 (first entry)
XX DE DNA encoding linker polypeptide for SCD40-GM-CSF fusion protein.
XX KW Systemic immune response; soluble form of CD40; sCD40; cytokine;
KW granulocyte-macrophage colony stimulating factor; GM-CSF; immunotherapy;
KW tumour; cancer; bacterial infection; viral infection;
KW transplant tissue rejection; autoimmune disease; gene therapy;
KW protein therapy; cytostatic; immunostimulant; ds; AIDS;
KW acquired immunodeficiency syndrome.
XX OS Synthetic.
XX XX WO200178769-A2.
XX PN 25-OCT-2001.
XX PD
XX PF 12-APR-2001; 2001WO-US012003.
XX PR 12-APR-2000; 2000US-0196489P.
XX PA (UYSF-) UNIV SOUTH FLORIDA.
XX XX Yu H, Turner JG, Tan J;
XX PI WPI; 2002-011074/01.
XX DR
XX PT Stimulating systemic immune response against diseases, e.g., cancer and
PT infection, comprises administering soluble CD40 optionally in combination
PT with a cytokine and/or cell based or isolated target antigen.
XX PS Disclosure; Page 8; 38pp; English.
XX CC The present invention relates to a method of stimulating a systemic
CC immune response for the treatment or prevention of animal and human
CC diseases. The method comprises administering a soluble form of CD40
CC (sCD40), optionally in combination with a cytokine and/or cell-based
CC antigen. Co-administration of sCD40 and granulocyte-macrophage colony
CC stimulating factor (GM-CSF) can be used as immunotherapy for the

```



CC treatment and prevention of tumours. The method of the present invention  
 CC is useful for stimulating a systemic immune response in a subject against  
 CC diseases such as cancer, bacterial or viral infection such as acquired  
 CC immunodeficiency syndrome (AIDS), transplant tissue rejection or  
 CC autoimmune disease. The method is also useful for treating or preventing  
 CC various diseases through the choice of antigen (specific for a condition)  
 CC and optionally co-administering it. The method is also useful for  
 CC suppressing growth of a tumour in a subject. The use of sCD40 gene  
 CC therapy for inducing anti-tumour immune response in vivo instead of  
 CC native CD40 gene therapy is more efficient as the soluble protein can  
 CC reach more effector cells than a membrane bound protein. The present  
 CC sequence encodes for a linker polypeptide used to link sCD40 to GM-CSF to  
 CC generate a fusion protein in the methods of the present invention  
 XX  
 SQ Sequence 12 BP; 0 A; 8 C; 4 G; 0 T; 0 U; 0 Other;

Query Match 100.0%; Score 9; DB 6; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 11 GCGGCGGCG 3

RESULT 13  
 AAL41283  
 ID AAL41283 standard; DNA; 12 BP.

XX AAL41283;

DT 02-JAN-2003 (first entry)

XX Oligonucleotide of DP/DF epitope.

DE Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
 KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
 KW allergic asthma; allergic rhinitis; atopic eczema; DP/DF; epitope; ds.

XX Dermatophagoides pteronyssinus.

OS WO200278736-A2.

PN 10-OCT-2002.

PF 28-MAR-2002; 2002WO-FR001098.

XX 30-MAR-2001; 2001FR-00004370.

PR 03-MAY-2001; 2001FR-00005929.

XX 29-MAY-2001; 2001US-00867159.

PA (ANTI-) ANTIALIS SARL.

XX Loria E, Terrasse G, Trehin Y;

DR WPI; 2002-750636/81.

XX Antiallergic compositions containing an anti-histamine, a histamine  
 PT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
 PT for the allergen.

XX Claim 20; Page 13; 32pp; French.

XX The invention relates to antiallergic compositions containing an anti-  
 CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
 CC isolated nucleic acid molecule that has at least one polynucleotide  
 CC sequence coding for the allergen, together with a pharmaceutical carrier.  
 CC The pharmaceutical composition of the invention is useful as a non-  
 CC specific antiallergic treatment, and also useful in the treatment of  
 CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
 CC allergic and atopic eczema. This polynucleotide sequence represents an  
 CC oligonucleotide of a DP/DF epitope relating to the antiallergic  
 CC compositions of the invention

XX Sequence 12 BP; 1 A; 3 C; 7 G; 1 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 9; DB 6; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 4 GCGGCGGCG 12

RESULT 14  
 ABQ80834  
 ID ABQ80834 standard; DNA; 12 BP.

XX ABQ80834;

DT 13-DEC-2002 (first entry)

DE Cysteine protease PCR primer.

XX Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
 KW immunotherapy; allergen; allergic hypersensitivity reaction;  
 KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
 KW cysteine protease; PCR; primer; ss.

OS Dermatophagoides pteronyssinus.

PN FR2822709-A1.

PD 04-OCT-2002.

PF 03-MAY-2001; 2001FR-00005929.

XX 30-MAR-2001; 2001FR-00004370.

PA (ANTI-) ANTIALIS SARL.

PI Loria E, Terrasse G, Trehin Y;

DR WPI; 2002-735037/80.

XX Antiallergic composition, useful for preventing and treating e.g. asthma,  
 PT rhinitis or eczema, containing at least two of allergen, antihistamine  
 PT and histamine synthesis inhibitor.

XX Claim 19; Page 7; 33pp; French.

XX The present invention relates to an antiallergic pharmaceutical  
 CC composition (I) comprising a pharmaceutical carrier containing an active  
 CC agent combination of at least two of: an allergen; an antihistamine; and  
 CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
 CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
 CC rhinitis or allergic atopic eczema, in babies, children or adults. To  
 CC illustrate the invention, cysteine protease from Dermatophagoides  
 CC pteronyssinus (AB98533), was used as an allergen. The present sequence  
 CC is a PCR primer, used to illustrate the invention

XX Sequence 12 BP; 1 A; 3 C; 7 G; 1 T; 0 U; 0 Other;

Query Match 100.0%; Score 9; DB 6; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
 |||||  
 Db 4 GCGGCGGCG 12

RESULT 15  
 ABX80021  
 ID ABX80021 standard; cDNA; 12 BP.

```

XX AC AX80021;
XX DT 17-APR-2003 (first entry)
XX DE EST polymorphic DNA repeat polynucleotide #346.
XX KW EST; expressed sequence tag; ss; polymorphic repeat; tandem repeat;
XX KW polymorphic marker prediction of ubiquitous simple sequences; POMPOUS;
XX KW Rep-X; human; genetic disease; drug-treatment; Machado-Joseph;
XX KW Haw River syndrome; Huntington's disease; fragile-X syndrome;
XX KW Friedrich's ataxia; myotonic dystrophy; hyperandrogenaemia;
XX KW spinal atrophy; bulbar atrophy; spinocerebellar ataxia.
XX OS Homo sapiens.
XX PN US6472154-B1.
XX PD 29-OCT-2002.
XX PF 31-DEC-1999; 99US-00475947.
XX PR 31-DEC-1999; 99US-00475947.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PI Garner HR, Wren JD, Minna JD, Fondon JW;
XX DR WPI; 2003-208818/20.
XX PT Identifying a candidate polymorphic repeat within a coding sequence, for
PT understanding or treating genetic disease, comprises detecting tandem
PT repeats in a target coding sequence and scoring the repeats for
PT polymorphic probability.
XX PS Example; Col 1171; 588pp; English.
XX CC The invention discloses a method for identifying a candidate polymorphic
CC repeat within a coding sequence (expressed sequence tag, EST), which
CC comprises detecting tandem repeats in a target coding sequence, scoring
CC the repeats for polymorphic probability and generating a dataset
CC correlating the repeats with polymorphic probability to identify a
CC candidate polymorphic repeat. The computational methods (polymorphic
CC marker prediction of ubiquitous simple sequences, POMPOUS, and Rep-X) are
CC useful for identifying and detecting candidate polymorphic repeats in
CC human genes, which can be used to understand, treat or eliminate genetic
CC diseases, predispositions or adverse drug-treatment reactions. Examples
CC of diseases linked to nucleotide repeats are Machado-Joseph, Haw River
CC syndrome, Huntington's disease, fragile-X syndrome, Friedrich's ataxia,
CC myotonic dystrophy, hyperandrogenaemia, spinal and bulbar atrophy and
CC spinocerebellar ataxia. The sequences presented in AX80022 are
CC the polymorphic repeats identified for a search of human ESTs
XX SQ Sequence 12 BP; 1 A; 4 C; 7 G; 0 T; 0 U; 0 Other;
Query Match 100.0%; Score 9; DB 8; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGGCGGCG 9
Db 1 GCGGCGGCG 9

```

Search completed: October 28, 2004, 22:43:34  
 Job time : 16.1616 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 23:44:08 ; Search time 82.559 Seconds  
(without alignments)  
558.975 Million cell updates/sec

Title: US-09-867-159A-6

Perfect score: 9

Sequence: 1 gcgggcgcg 9

Scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	100.0	9	10	US-09-867-159A-6
2	9	100.0	10	16	US-10-301-875A-32
3	9	100.0	10	18	US-10-670-011-388
4	9	100.0	12	9	US-09-832-865-1
5	9	100.0	12	10	US-09-867-159A-7
6	9	100.0	12	15	US-10-091-281-457
7	9	100.0	15	14	US-10-132-002-13
8	9	100.0	15	15	US-10-339-674-1508
9	9	100.0	15	15	US-10-339-674-1625
10	9	100.0	16	15	US-10-339-674-561
11	9	100.0	16	15	US-10-339-674-1288
12	9	100.0	16	16	US-10-138-674-5650

13 9 100.0 16 17 US-10-287-949A-5650  
14 9 100.0 17 9 US-09-227-742-11  
15 9 100.0 17 10 US-09-780-533A-1787  
16 9 100.0 17 10 US-09-780-533A-1788  
17 9 100.0 17 10 US-09-780-533A-1789  
18 9 100.0 17 10 US-09-780-533A-2336  
19 9 100.0 17 10 US-09-780-533A-2337  
20 9 100.0 17 10 US-09-780-533A-2338  
21 9 100.0 17 10 US-09-848-754A-831  
22 9 100.0 17 10 US-09-848-754A-832  
23 9 100.0 17 10 US-09-848-754A-833  
24 9 100.0 17 10 US-09-848-754A-2124  
25 9 100.0 17 10 US-09-848-754A-2125  
26 9 100.0 17 10 US-09-848-754A-2126  
27 9 100.0 17 10 US-09-848-754A-2127  
28 9 100.0 17 10 US-09-930-423-313  
29 9 100.0 17 10 US-09-930-423-314  
30 9 100.0 17 10 US-09-930-423-972  
31 9 100.0 17 10 US-09-930-423-973  
32 9 100.0 17 10 US-09-930-423-974  
33 9 100.0 17 10 US-09-745-237A-313  
34 9 100.0 17 10 US-09-745-237A-314  
35 9 100.0 17 10 US-09-745-237A-972  
36 9 100.0 17 10 US-09-745-237A-973  
37 9 100.0 17 10 US-09-745-237A-974  
38 9 100.0 17 14 US-10-138-970A-10  
39 9 100.0 17 15 US-10-156-306-2261  
40 9 100.0 17 15 US-10-156-306-2433  
41 9 100.0 17 15 US-10-156-306-2434  
42 9 100.0 17 15 US-10-156-306-3486  
43 9 100.0 17 15 US-10-156-306-3487  
44 9 100.0 17 15 US-10-238-700-3  
45 9 100.0 17 15 US-10-238-700-4

#### ALIGNMENTS

#### RESULT 1

US-09-867-159A-6  
; Sequence 6, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 9  
; TYPE: DNA  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: primer  
; LOCATION: (1)..(9)  
; OTHER INFORMATION:  
US-09-867-159A-6

Query Match 100.0%; Score 9; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 5.7e+08;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9

DB 1 GCGGCGGCG 9

TERRASSE, GAETAN

LORIA, EMILE

TREHIN, YVES

```
RESULT 2
US-10-301-875A-32/c
; Sequence 32, Application US/10301875A
; Publication No. US20040091874A1
; GENERAL INFORMATION:
; APPLICANT: YAMASAKI, KAZUHIKO
; TITLE OF INVENTION: SENSOR CHIP FOR NUCLEIC ACID SELECTION
; FILE REFERENCE: 081356/0179
; CURRENT APPLICATION NUMBER: US/10/301.875A
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: JP 2002/149330
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
US-10-301-875A-32

Query Match      100.0%; Score 9; DB 16; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.9e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGCGGCG 9
        |||||
DB      1 GCGGCGGCG 9
        |||||

RESULT 3
US-10-670-011-388
; Sequence 388, Application US/10670011
; Publication No. US20040209832A1
; GENERAL INFORMATION:
; APPLICANT: Sirna Therapeutics, Inc.
; APPLICANT: McSwiggen, James
; APPLICANT: Beigelman, Leonid
; APPLICANT: Pavco, Pamela
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Vascular Endothelial
; TITLE OF INVENTION: Growth Factor and Vascular Endothelial Growth Factor Receptor
; TITLE OF INVENTION: Gene Expression Using Short Interfering Nucleic Acid (siNA)
; FILE REFERENCE: 400/132 (MEHB02-742-G)
; CURRENT APPLICATION NUMBER: US/10/670,011
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: PCT/US03/05022
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US60/358,580
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US60/363,124
; PRIOR FILING DATE: 2002-03-11
; PRIOR APPLICATION NUMBER: US60/386,782
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US60/393,796
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US60/399,348
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US60/406,784
; PRIOR FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: US60/408,378
; PRIOR FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US60/409,293
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: US60/440,129
; PRIOR FILING DATE: 2003-01-15
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 427
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 388
        1 GCGGCGGCG 9
        |||||
        9 GCGGCGGCG 1
        |||||

RESULT 4
US-09-832-865-1/c
; Sequence 1, Application US/09832865
; Publication No. US2002022017A1
; GENERAL INFORMATION:
; APPLICANT: Hua, Yu
; TITLE OF INVENTION: Regulation of Systemic Immune Responses Utilizing Soluble CD40
; FILE REFERENCE: USF-T142X
; CURRENT APPLICATION NUMBER: US/09/832.865
; CURRENT FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: US 60/196,489
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic nucleotide sequence encoding polypeptide linker
US-09-832-865-1

Query Match      100.0%; Score 9; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGCGGCG 9
        |||||
DB      1 GCGGCGGCG 9
        |||||

RESULT 5
US-09-867-159A-7
; Sequence 7, Application US/09867159A
; Publication No. US20030104013A1
; GENERAL INFORMATION:
; APPLICANT: ANTIALIS
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one
; TITLE OF INVENTION: and at least one anti-histamine compound
; FILE REFERENCE: B112812US-antialis
; CURRENT APPLICATION NUMBER: US/09/867,159A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: FR01/04370
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: FR01/05929
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: Primer
; LOCATION: (1)..(12)
; OTHER INFORMATION:
US-09-867-159A-7

Query Match      100.0%; Score 9; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGCGGCG 9
        |||||
DB      1 GCGGCGGCG 3
        |||||

RESULT 6
US-09-867-159A-7
; Sequence 7, Application US/09867159A
; Publication No. US20030104013A1
; GENERAL INFORMATION:
; APPLICANT: ANTIALIS
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one
; TITLE OF INVENTION: and at least one anti-histamine compound
; FILE REFERENCE: B112812US-antialis
; CURRENT APPLICATION NUMBER: US/09/867,159A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: FR01/04370
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: FR01/05929
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Dermatophagoides pteronyssinus
; FEATURE:
; NAME/KEY: Primer
; LOCATION: (1)..(12)
; OTHER INFORMATION:
US-09-867-159A-7
```

```
Query Match      100.0%; Score 9; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5e+04; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGCGCG 9
DB 4 GCGCGCGCG 12

RESULT 6
US-10-091-281-457
; Sequence 457, Application US/10091281
; Publication No. US20030190617A1
; GENERAL INFORMATION:
; APPLICANT: RAYMOND, VINCENT
; APPLICANT: SI, ERWIN
; APPLICANT: MORISSETTE, JEAN
; TITLE OF INVENTION: OPTINEURIN NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: 13587.338
; CURRENT APPLICATION NUMBER: US/10/091,281
; CURRENT FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: PatentIn Ver: 2.1
; SEQ ID NO 457
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Putative EGRF/NGFIC.01 motif
US-10-091-281-457

Query Match      100.0%; Score 9; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.5e+04; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGCGCG 9
DB 3 GCGCGCGCG 11

RESULT 7
US-10-132-002-13
; Sequence 13, Application US/10132002
; Publication No. US20030022204A1
; GENERAL INFORMATION:
; APPLICANT: Lansdorp, Peter
; TITLE OF INVENTION: Method for Detecting Multiple Copies of
; a Repeat Sequence in a Nucleic Acid Molecule
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWSON & HOWSON
; STREET: 321 No. US20030022204A1ristown Road
; CITY: Spring House
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/132,002
; FILING DATE: 25-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/730,635
; FILING DATE: 11-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: B&P7USA
```

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9200
; TELEFAX: (215) 540-5818
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (Genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-132-002-13

Query Match      100.0%; Score 9; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.1e+04; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
DB 2 GCGGCGGCG 10

RESULT 8
US-10-339-674-1508/c
; Sequence 1508, Application US/10339674
; Publication No. US20030204318A1
; GENERAL INFORMATION:
; APPLICANT: Feldmann, Richard J.; Global Determinants, Inc.
; TITLE OF INVENTION: Escherichia coli K-12 MG1655 complete genome.
; FILE REFERENCE: Jim Zegeer Law Offices - 703-684-8333
; CURRENT APPLICATION NUMBER: US/10/339,674
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 3537
; SOFTWARE: Proprietary
; SEQ ID NO 1508
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Escherichia coli K-12 MG1655 complete genome.
; FEATURE:
; LOCATION: (2089336)...(2089350)
; OTHER INFORMATION: Chromosome = 1 Strand = positive
US-10-339-674-1508

Query Match      100.0%; Score 9; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.1e+04; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
DB 10 GCGGCGGCG 2

RESULT 9
US-10-339-674-1625/c
; Sequence 1625, Application US/10339674
; Publication No. US20030204318A1
; GENERAL INFORMATION:
; APPLICANT: Feldmann, Richard J.; Global Determinants, Inc.
; TITLE OF INVENTION: Escherichia coli K-12 MG1655 complete genome.
; FILE REFERENCE: Jim Zegeer Law Offices - 703-684-8333
; CURRENT APPLICATION NUMBER: US/10/339,674
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 3537
; SOFTWARE: Proprietary
; SEQ ID NO 1625
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Escherichia coli K-12 MG1655 complete genome.
; FEATURE:
; LOCATION: (2250971)...(2250984)
; OTHER INFORMATION: Chromosome = 1 Strand = positive
US-10-339-674-1625
```

Query Match 100.0%; Score 9; DB 15; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 Db 10 GCGGCGGCG 2

RESULT 10  
 US-10-339-674-561/c  
 ; Sequence 561, Application US/10339674  
 ; Publication No. US20030204318A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Feldmann, Richard J.; Global Determinants, Inc.  
 ; TITLE OF INVENTION: Escherichia coli K-12 MG1655 complete genome.  
 ; FILE REFERENCE: Jim Zeeger Law Offices - 703-684-8333  
 ; CURRENT APPLICATION NUMBER: US/10339,674  
 ; CURRENT FILING DATE: 2003-06-06  
 ; NUMBER OF SEQ ID NOS: 3537  
 ; SOFTWARE: Proprietary  
 ; SEQ ID NO 561  
 ; LENGTH: 16  
 ; TYPE: DNA  
 ; ORGANISM: Escherichia coli K-12 MG1655 complete genome.  
 ; FEATURE:  
 ; LOCATION: (623181)...(623196)  
 ; OTHER INFORMATION: Chromosome = 1 Strand = positive ConnectronObjectNumber = 748  
 US-10-339-674-561

Query Match 100.0%; Score 9; DB 15; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 Db 13 GCGGCGGCG 5

RESULT 11  
 US-10-339-674-1288/c  
 ; Sequence 1288, Application US/10339674  
 ; Publication No. US20030204318A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Feldmann, Richard J.; Global Determinants, Inc.  
 ; TITLE OF INVENTION: Escherichia coli K-12 MG1655 complete genome.  
 ; FILE REFERENCE: Jim Zeeger Law Offices - 703-684-8333  
 ; CURRENT APPLICATION NUMBER: US/10339,674  
 ; CURRENT FILING DATE: 2003-06-06  
 ; NUMBER OF SEQ ID NOS: 3537  
 ; SOFTWARE: Proprietary  
 ; SEQ ID NO 1288  
 ; LENGTH: 16  
 ; TYPE: DNA  
 ; ORGANISM: Escherichia coli K-12 MG1655 complete genome.  
 ; FEATURE:  
 ; LOCATION: (1548200)...(1548215)  
 ; OTHER INFORMATION: Chromosome = 1 Strand = negative ConnectronObjectNumber = 1693  
 US-10-339-674-1288

Query Match 100.0%; Score 9; DB 15; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 Db 13 GCGGCGGCG 5

RESULT 12  
 US-10-138-674-5650  
 ; Sequence 5650, Application US/10138674

Publication No. US20040077565A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
 ; APPLICANT: Pavco, Pam  
 ; APPLICANT: McSwiggen, Jim  
 ; APPLICANT: Stinchcomb, Dan  
 ; APPLICANT: Escobedo, Jaime  
 ; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Rel  
 ; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
 ; FILE REFERENCE: MBH00-876-N (400/049)  
 ; CURRENT APPLICATION NUMBER: US/10/138,674  
 ; CURRENT FILING DATE: 2002-05-03  
 ; NUMBER OF SEQ ID NOS: 20822  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 5650  
 ; LENGTH: 16  
 ; TYPE: RNA  
 ; ORGANISM: Homo sapiens  
 US-10-138-674-5650

Query Match 100.0%; Score 9; DB 16; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 Db 3 GCGGCGGCG 11

RESULT 13  
 US-10-287-949A-5650  
 ; Sequence 5650, Application US/10287949A  
 ; Publication No. US20040102389A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
 ; APPLICANT: Pavco, Pam  
 ; APPLICANT: McSwiggen, Jim  
 ; APPLICANT: Stinchcomb, Dan  
 ; APPLICANT: Escobedo, Jaime  
 ; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Rel  
 ; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
 ; FILE REFERENCE: MBH00-876-N (400/049)  
 ; CURRENT APPLICATION NUMBER: US/10/287,949A  
 ; CURRENT FILING DATE: 2003-04-11  
 ; NUMBER OF SEQ ID NOS: 20822  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 5650  
 ; LENGTH: 16  
 ; TYPE: RNA  
 ; ORGANISM: Homo sapiens  
 US-10-287-949A-5650

Query Match 100.0%; Score 9; DB 17; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+04;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
 Db 3 GCGGCGGCG 11

RESULT 14  
 US-09-227-742-11  
 ; Sequence 11, Application US/09227742  
 ; Patent No. US20020137191A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BLOOM, FREDERIC  
 ; APPLICANT: KUO, JONATHAN  
 ; APPLICANT: LIN, JHY-JHU  
 ; APPLICANT: MA, JIN  
 ; TITLE OF INVENTION: METHOD FOR INCREASING VIABILITY  
 ; TITLE OF INVENTION: AND TRANSFORMATION EFFICIENCY OF BACTERIA DURING  
 ; TITLE OF INVENTION: STORAGE AT LOW TEMPERATURES

NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howrey & Simon  
STREET: 1299 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004-2402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/227,742  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/826,426  
FILING DATE: 27-MAR-1997  
APPLICATION NUMBER: U.S. Ser. No. US20020137191A1 60/025,838  
FILING DATE: 05-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Auerbach, Jeffrey  
REGISTRATION NUMBER: 32,680  
REFERENCE/DOCKET NUMBER: 04227-0031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 383-7451  
TELEFAX: 202 383-6610  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-227-742-11

Query Match 100.0%; Score 9; DB 9; Length 17;  
Best Local Similarity 100.0%; Pred. No. 6.8e+04;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGCG 9  
Db 4 GCGGCGGCG 12

## RESULT 15

US-09-780-533A-1787  
Sequence 1787, Application US/09780533A  
Publication No. US20030060611A1  
GENERAL INFORMATION:  
APPLICANT: Ribozyme Pharmaceuticals, Inc.  
APPLICANT: Blatt, Larry  
APPLICANT: McSwiggen, Jim  
APPLICANT: Chowkira, Bharat  
APPLICANT: Haerberli, Pete  
TITLE OF INVENTION: Method and Reagent for the Inhibition of NCO Gene  
FILE REFERENCE: MEH800,878-A (400/011)  
CURRENT APPLICATION NUMBER: US/09/780,533A  
CURRENT FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: US 60/181,797  
PRIOR FILING DATE: 2000-02-11  
NUMBER OF SEQ ID NOS: 6679  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 1787  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Homo sapiens  
US-09-780-533A-1787

Query Match 100.0%; Score 9; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 6.8e+04;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCGGCGGCG 9  
Db 9 GCGGCGGCG 17

Search completed: October 29, 2004, 06:26:13  
Job time : 84.559 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 22:29:43 ; Search time 2.46288 Seconds  
(without alignments)  
2597.404 Million cell updates/sec

Title: US-09-867-159A-6

Perfect score: 9

Sequence: 1 GCGGCGGCG 9

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
  - 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
  - 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
  - 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
  - 5: /cgn2\_6/prodata/1/ina/PCUTS\_COMB.seq:\*
  - 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	100.0	12	2	US-08-457-273B-38
2	9	100.0	12	2	US-08-863-639A-95
3	9	100.0	12	4	US-09-475-947A-346
4	9	100.0	12	4	US-09-724-028A-22
C 5	9	100.0	14	3	US-08-985-162-1759
C 6	9	100.0	14	3	US-08-985-162-1759
C 7	9	100.0	14	4	US-09-401-063-1759
8	9	100.0	14	4	US-09-401-063-1759
C 9	9	100.0	15	2	US-08-863-639A-21
10	9	100.0	15	4	US-08-730-635-13
11	9	100.0	16	4	US-09-371-772B-5650
12	9	100.0	16	4	US-09-479-005A-2
13	9	100.0	17	2	US-08-826-426-11
C 14	9	100.0	17	3	US-08-909-742-3
C 15	9	100.0	17	3	US-08-909-742-3
C 16	9	100.0	17	3	US-09-412-289-3
C 17	9	100.0	17	3	US-09-412-289-4
18	9	100.0	17	4	US-09-343-698-1
C 19	9	100.0	17	4	US-09-343-698-2
20	9	100.0	17	4	US-08-325-955-1
C 21	9	100.0	17	4	US-08-325-955-2
C 22	9	100.0	18	2	US-09-205-860-10
C 23	9	100.0	18	2	US-09-205-860-11
24	9	100.0	18	2	US-08-857-946-8
25	9	100.0	18	2	US-08-857-946-14
C 26	9	100.0	18	2	US-08-857-946-15
27	9	100.0	18	3	US-09-205-921-8

28	9	100.0	18	3	US-09-205-921-9	Sequence 9, Appli
29	9	100.0	18	3	US-08-970-740-8	Sequence 8, Appli
30	9	100.0	18	3	US-08-970-740-14	Sequence 14, Appli
C 31	9	100.0	18	3	US-08-970-740-15	Sequence 15, Appli
C 32	9	100.0	18	3	US-09-358-381-13	Sequence 13, Appli
33	9	100.0	18	3	US-09-143-212-44	Sequence 44, Appli
34	9	100.0	18	3	US-09-143-212-45	Sequence 45, Appli
35	9	100.0	18	3	US-09-143-212-45	Sequence 46, Appli
C 36	9	100.0	18	3	US-09-496-694B-99	Sequence 99, Appli
C 37	9	100.0	18	3	US-08-679-645-1163	Sequence 1163, Ap
C 38	9	100.0	18	3	US-08-679-645-1165	Sequence 1165, Ap
C 39	9	100.0	18	3	US-08-679-645-1167	Sequence 1167, Ap
C 40	9	100.0	18	3	US-08-679-645-1169	Sequence 1169, Ap
C 41	9	100.0	18	3	US-08-679-645-1169	Sequence 25, Appli
C 42	9	100.0	18	4	US-09-500-700-68	Sequence 68, Appli
43	9	100.0	19	2	US-08-900-148-3	Sequence 3, Appli
C 44	9	100.0	20	2	US-08-465-485A-28	Sequence 28, Appli
45	9	100.0	20	2	US-08-465-485A-28	

ALIGNMENTS

RESULT 1  
US-08-457-273B-38/c  
; Sequence 38, Application US/08457273B  
; Patent No. 5849995  
; GENERAL INFORMATION:  
; APPLICANT: Hayden, Michael  
; APPLICANT: Lin, Biacyang  
; APPLICANT: Nasir, Jamal  
; TITLE OF INVENTION: Mouse Model for Huntington's Disease and  
; TITLE OF INVENTION: Related DNA Sequences  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Virginia Bennett  
; STREET: PO Box 37428  
; CITY: Raleigh  
; STATE: No. 5849995th Carolina  
; COUNTRY: US  
; ZIP: 27627  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/457,273B  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bennett, Virginia C.  
; REGISTRATION NUMBER: 37,092  
; REFERENCE/DOCKET NUMBER: 3477-85A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-854-1400  
; TELEFAX: 919-854-1401  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-457-273B-38

Query Match 100.0%; Score 9; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCGGCG 9  
DB 11 GCGGCGGCG 3

```

; US-08-863-639A-95
; Sequence 95, Application US/08863639A
; Patent No. 5981185
; GENERAL INFORMATION:
; APPLICANT: Matson, Robert S.
; APPLICANT: Coassin, Peter J.
; APPLICANT: Rampal, Jang B.
; APPLICANT: Caskey, C. T.
; TITLE OF INVENTION: OLIGONUCLEOTIDE REPEAT ARRAYS
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, 9th Floor
; CITY: Pasadena
; STATE: CA
; COUNTRY: USA
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Corel WordPerfect 8 version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/863,639A
; FILING DATE: May 28, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Joseph E. Muech
; REGISTRATION NUMBER: 20,532
; REFERENCE/DOCKET NUMBER: 11859-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (626) 796-4000
; TELEFAX: (626) 795-6321
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other nucleic acid
; US-08-863-639A-95

Query Match 100.0%; Score 9; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
Db 1 GCGGCGGCG 9

RESULT 3
US-09-475-947A-346
; Sequence 346, Application US/09475947A
; Patent No. 6472154
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTSD0667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 346
; LENGTH: 12
; TYPE: DNA
; ORGANISM: human
; US-09-475-947A-346

Query Match 100.0%; Score 9; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
Db 1 GCGGCGGCG 9

RESULT 4
US-09-724-028A-22
; Sequence 22, Application US/09724028A
; Patent No. 6777180
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; TITLE OF INVENTION: METHOD FOR FULL-LENGTH cDNA CLONING
; TITLE OF INVENTION: USING DEGENERATE STEM LOOP ANNEALING PRIMERS
; FILE REFERENCE: A34701 (070050.1728)
; CURRENT APPLICATION NUMBER: US/09/724,028A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 12
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the 5' end of the C-ORF product of
; Patent No. 6777180
; OTHER INFORMATION: mda-9
; US-09-724-028A-22

Query Match 100.0%; Score 9; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
Db 2 GCGGCGGCG 10

RESULT 5
US-08-985-162-1759/c
; Sequence 1759, Application US/08985162
; Patent No. 6057156
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir
; APPLICANT: Fell, Patricia
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; TITLE OF INVENTION: FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,162
; FILING DATE: 04 December 1997
; US-08-985-162-1759/c
```

```
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/036,476
; FILING DATE: 31 January 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1759:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-162-1759

Query Match 100.0%; Score 9; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
Db 13 GCGGCGGCG 5

RESULT 6
US-08-985-162-1760/c
; Sequence 1760, Application US/08985162
; Patent No. 6057156
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir
; APPLICANT: Fell, Patricia
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMAIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; TITLE OF INVENTION: FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,162
; FILING DATE: 31 January 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/036,476
; FILING DATE: 04 December 1997
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1760:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-162-1760
```

```
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-162-1760

Query Match 100.0%; Score 9; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
Db 10 GCGGCGGCG 2

RESULT 7
US-09-401-063-1759/C
; Sequence 1759, Application US/09401063
; Patent No. 6623962
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir
; APPLICANT: Fell, Patricia
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMAIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; TITLE OF INVENTION: FACTOR RECEPTORS
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/401,063
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/985,162
; FILING DATE: 04 December 1997
; APPLICATION NUMBER: 60/036,476
; FILING DATE: 31 January 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1759:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-401-063-1759

Query Match 100.0%; Score 9; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9
```



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 540-9200  
TELEFAX: (215) 540-5818  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-730-635-13

Query Match 100.0%; Score 9; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 3.2e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
Db 2 GCGGCGGCG 10

RESULT 11  
US-09-371-772B-5650  
Sequence 5650, Application US/09371772B  
Patent No. 6566127  
GENERAL INFORMATION:  
APPLICANT: Ribozyme Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwigger, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Related to the Growth of Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MEHB00-876-J (237/198)  
CURRENT APPLICATION NUMBER: US/09/371,772B  
CURRENT FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 14225  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5650  
LENGTH: 16  
TYPE: RNA  
ORGANISM: Homo sapiens  
US-09-371-772B-5650

Query Match 100.0%; Score 9; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.1e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
Db 3 GCGGCGGCG 11

RESULT 12  
US-09-479-005A-2  
Sequence 2, Application US/09479005A  
Patent No. 6656731  
GENERAL INFORMATION:  
APPLICANT: Ribozyme Pharmaceuticals, Inc.  
TITLE OF INVENTION: Nucleic Acid Catalysts with Endonuclease Activity  
FILE REFERENCE: MEHB00-884-C  
CURRENT APPLICATION NUMBER: US/09/479,005A  
CURRENT FILING DATE: 2000-01-07  
PRIOR APPLICATION NUMBER: US 09/444,209  
PRIOR FILING DATE: 1999-11-19  
PRIOR APPLICATION NUMBER: US 09/159,274  
PRIOR FILING DATE: 1998-09-22  
PRIOR APPLICATION NUMBER: US 60/059,473

PRIOR FILING DATE: 1997-09-22  
NUMBER OF SEQ ID NOS: 1208  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 16  
TYPE: RNA  
ORGANISM: Homo sapiens  
US-09-479-005A-2

Query Match 100.0%; Score 9; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.1e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9  
Db 3 GCGGCGGCG 11

RESULT 13  
US-08-826-426-11  
Sequence 11, Application US/08826426  
Patent No. 5891692  
GENERAL INFORMATION:  
APPLICANT: BLOOM, FREDERIC  
APPLICANT: KUO, JONATHAN  
APPLICANT: LIN, JHY-JHU  
APPLICANT: MA, JIN  
TITLE OF INVENTION: METHOD FOR INCREASING VIABILITY AND TRANSFORMATION EFFICIENCY OF BACTERIA DURING STORAGE AT LOW TEMPERATURES  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howrey & Simon  
STREET: 1299 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004-2402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Fast-Seq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/826,426  
FILING DATE: 27-VAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S. Ser. No. 5891692 60/014,330  
FILING DATE: 29-VAR-1996  
APPLICATION NUMBER: U.S. Ser. No. 5891692 60/025,838  
FILING DATE: 05-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Auerbach, Jeffrey  
REGISTRATION NUMBER: 32,680  
REFERENCE/DOCKET NUMBER: 04227-0031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 383-7451  
TELEFAX: 202 383-6610  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-826-426-11

Query Match 100.0%; Score 9; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.1e+04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGCG 9

```

DB      4 GCGGGCGG 12
RESULT 14
US-08-909-742-3/c
; Sequence 3, Application US/08909742
; Patent No. 6007991
; GENERAL INFORMATION:
; APPLICANT: Vimala S. Sivaraman
; APPLICANT: Hsien-Yu Wang
; APPLICANT: Craig C. Malbon
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES FOR MITOGEN-
; TITLE OF INVENTION: ACTIVATED PROTEIN KINASES AS THERAPY FOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for windows
; CURRENT APPLICATION DATA:
; FILING DATE: August 12, 1997
; APPLICATION NUMBER: US/08/909,742
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,994
; FILING DATE: April 1, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/827,520
; FILING DATE: March 28, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Adams, Lindsay S.
; REGISTRATION NUMBER: 36,425
; REFERENCE/DOCKET NUMBER: 178-225 CIP II
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-909-742-3
Query Match 100.0%; Score 9; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.le+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGGCGG 9
DB      14 GCGGGCGG 6
RESULT 15
US-08-909-742-4/c
; Sequence 4, Application US/08909742
; Patent No. 6007991
; GENERAL INFORMATION:
; APPLICANT: Vimala S. Sivaraman
; APPLICANT: Hsien-Yu Wang
; APPLICANT: Craig C. Malbon
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES FOR MITOGEN-
; TITLE OF INVENTION: ACTIVATED PROTEIN KINASES AS THERAPY FOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for windows
; CURRENT APPLICATION DATA:
; FILING DATE: August 12, 1997
; APPLICATION NUMBER: US/08/909,742
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/831,994
; FILING DATE: April 1, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/827,520
; FILING DATE: March 28, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Adams, Lindsay S.
; REGISTRATION NUMBER: 36,425
; REFERENCE/DOCKET NUMBER: 178-225 CIP II
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-909-742-4
Query Match 100.0%; Score 9; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.le+04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGGGCGG 9
DB      14 GCGGGCGG 6
Search completed: October 29, 2004, 01:48:22
Job time : 3.46288 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 20:26:03 ; Search time 63.1441 Seconds  
(without alignments)  
8987.005 Million cell updates/sec

Title: US-09-867-159A-7  
Perfect score: 12  
Sequence: 1 tgagcgccgccc 12

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	12	100.0	12	6	AX586978
2	12	100.0	51	6	AX915864
3	12	100.0	51	6	BD051397
4	12	100.0	80	9	HSTAL1D
5	12	100.0	124	10	AY598730
6	12	100.0	138	3	AF275765
7	12	100.0	123	6	E15662
8	12	100.0	182	9	HS37H12R
9	12	100.0	177	9	HUMSCL1
10	12	100.0	181	11	BV136875
11	12	100.0	183	11	BV136872
12	12	100.0	184	6	CO657470
13	12	100.0	184	11	BV136873
14	12	100.0	186	11	BV136865
15	12	100.0	193	11	BV136855
16	12	100.0	193	11	BV136867
17	12	100.0	202	11	BV136860
18	12	100.0	210	11	BV136862
19	12	100.0	216	6	AX205133

C 20	12	100.0	216	11	BV136866
C 21	12	100.0	219	6	AX119209
C 22	12	100.0	222	11	BV136858
C 23	12	100.0	224	8	AY023140
C 24	12	100.0	224	8	AY023172
C 25	12	100.0	224	8	AY023179
C 26	12	100.0	224	8	AY023179
C 27	12	100.0	234	11	BV136863
C 28	12	100.0	236	6	AX896897
C 29	12	100.0	236	6	BD032430
C 30	12	100.0	244	11	BV136869
C 31	12	100.0	248	11	BV136871
C 32	12	100.0	248	11	BV136874
C 33	12	100.0	249	11	BV136868
C 34	12	100.0	250	11	BV136857
C 35	12	100.0	250	11	BV136870
C 36	12	100.0	251	11	BV136861
C 37	12	100.0	256	11	BV136856
C 38	12	100.0	256	11	BV136859
C 39	12	100.0	256	11	BV136864
C 40	12	100.0	285	6	AR245278
C 41	12	100.0	286	6	AR251223
C 42	12	100.0	286	11	BV138471
C 43	12	100.0	287	11	BV138489
C 44	12	100.0	290	11	BV138496
C 45	12	100.0	291	6	CQ749383

#### ALIGNMENTS

AX586978 12 bp DNA linear PAT 10-JAN-2003  
Sequence 7 from Patent WO2078736.  
AX586978  
AX586978.1 GI:27655853  
Dermatophagoides pteronyssinus (European house dust mite)  
Dermatophagoides pteronyssinus  
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
Acariiformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;  
Pyroglyphidae; Dermatophagoides.  
1  
Loxia E., Terrasse G. and Trehin Y.  
Antiallergic pharmaceutical composition  
Patent: WO 02078736-A 7 10-OCT-2002;  
Antialis (FR)  
Location/Qualifiers  
1..12  
/organism="Dermatophagoides pteronyssinus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6956"

#### ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCAGCGCGCGCG 12  
DB 1 TCAGCGCGCGCG 12  
RESULT 2  
LOCUS AX915864 51 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 31727 from Patent EP1033401.  
ACCESSION AX915864  
VERSION AX915864.1 GI:40209653  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 31727 06-SEP-2000;
Genset (FR)
FEATURES
source Location/Qualifiers
1..51
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 12; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGAGCGCGCGCG 12
Db 33 TGAGCGCGCGCG 22
RESULT 3
LOCUS BD051397 51 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD051397
VERSION BD051397.1 GI:22593139
KEYWORDS JP 2001269182-A/27643.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 51)
REFERENCE
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 27643 02-OCT-2001;
GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001269182-A/27643
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
PI JORDAN
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC
C12N5/10.
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC
G06F15/40
CC
FH Key Location/Qualifiers.
FEATURES
source Location/Qualifiers
1..51
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 12; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGAGCGCGCGCG 12
Db 33 TGAGCGCGCGCG 22
RESULT 4
HSTALID/c
LOCUS HSTALID Human tal-1 gene, exon 1a.
DEFINITION
ACCESSION X58621

```

```

X58621.1 GI:36683
Proto-oncogene; TAL1 gene.
Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 80)
REFERENCE
AUTHORS Chen,Q., Yang,C.Y., Tsan,J.T., Xia,Y., Ragab,A.H., Peiper,S.C.,
Carroll,A. and Baer,R.
TITLE Coding sequences of the tal-1 gene are disrupted by chromosome
translocation in human T cell leukemia
JOURNAL J. Exp. Med. 172 (5), 1403-1408 (1990)
MEDLINE 91037802
PUBMED 2230650
FEATURES
source Location/Qualifiers
1..80
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..71
/organism="tal-1"
1..71
/organism="tal-1"
/organism="tal-1"
/codon_start=1
/product="tal-1 product"
/protein_id="CAA41476.1"
/db_xref="GI:36684"
/db_xref="GOA:PI7542"
/db_xref="Swiss-Prot:PI7542"
/translation="RGRRAAAQDRASKWPHAYPRSGXT"
1..71
/organism="tal-1"
/feature="exon 1a"
/number=1
72..80
ORIGIN
Query Match 100.0%; Score 12; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGAGCGCGCGCG 12
Db 23 TGAGCGCGCGCG 12
RESULT 5
AV598730/c
LOCUS AV598730 124 bp mRNA linear ROD 18-MAY-2004
DEFINITION Cricetulus griseus N-acetylglucosaminyltransferase III mutant
LEC10A (Mgat3) mRNA, 5' UTR.
ACCESSION AV598730
VERSION AV598730.1 GI:47028354
KEYWORDS Cricetulus griseus (Chinese hamster)
SOURCE Cricetulus griseus
ORGANISM Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
1 (bases 1 to 124)
REFERENCE
AUTHORS Stanley,P., Sundaram,S., Tang,J. and Shi,S.
TITLE Molecular Analysis of Three Gain-of-Function CHO Mutants That Add
the Bisecting GlcNAc to N-Glycans
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 124)
AUTHORS Yang,X., Shi,S. and Stanley,P.
TITLE Direct Submission
JOURNAL Submitted (13-APR-2004) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA
FEATURES
source Location/Qualifiers
1..124
/organism="Cricetulus griseus"
/mol_type="mRNA"

```



```

/db_xref="taxon:10029"
/cell_line="CHO"
<1..>124
/genes="Mgat3"
/notes="N-acetylglucosaminyltransferase III mutant LEC10A"
<1..>124
/genes="Mgat3"

ORIGIN
Query Match 100.0%; Score 12; DB 10; Length 124;
Best Local Similarity 100.0%; Pred.No. 1e+05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
Db 50 TGAGCGCGCGCG 39

RESULT 6
AF275765 138 bp DNA linear INV 25-NOV-2001
LOCUS Epitedia wenmanni engrailed (en) gene, partial cds.
DEFINITION Epitedia wenmanni engrailed (en) gene, partial cds.
ACCESSION AF275765
VERSION AF275765.1 GI:17063311
KEYWORDS
SOURCE Epitedia wenmanni
ORGANISM Epitedia wenmanni
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Siphonaptera; Hystriochpsylloidea;
Ctenophthalmidae; Epitedia.
REFERENCE 1 (bases 1 to 138)
AUTHORS Whiting,M.F.
TITLE Multiple independent origins of the engrailed class homeobox intron
in insects: One character taxonomy, homoplasy, and the Strepsiptera
problem
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 138)
AUTHORS Whiting,M.F.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2000) Zoology, Brigham Young University, 574
Wildsee Building, Provo, UT 84602, USA
FEATURES
source
1..138
Location/Qualifiers
/organism="Epitedia wenmanni"
/mol_type="genomic DNA"
/db_xref="taxon:173793"
<1..>138
/genes="en"
/genes="en"
/product="engrailed"
<1..>138
/genes="en"
/codon_start=1
/product="engrailed"
/protein_id="AAL34996.1"
/db_xref="GI:17063312"
/translation="KPRPTAFSCAQLARLKHEAENRYLTERRRQQLSARLGLNEAQI
KI"

ORIGIN
Query Match 100.0%; Score 12; DB 3; Length 138;
Best Local Similarity 100.0%; Pred.No. 9.9e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
Db 78 TGAGCGCGCGCG 89

RESULT 7
E15662/c 153 bp DNA linear PAT 28-JUL-1999
LOCUS

```

```

DEFINITION Mouse mRNA for protamine 1.
ACCESSION E15662
VERSION E15662.1 GI:5710345
KEYWORDS JP 1998081698-A/9.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 153)
AUTHORS Takemura,F., Ueno,H. and Ito,S.
TITLE NUCLEIC ACID-BONDING POLYPEPTIDE, ITS PRODUCTION AND USE
JOURNAL Patent: JP 1998081698-A 9 31-MAR-1998;
FUJIREBIO INC
COMMENT
OS Mus sp. (mouse)
PN JP 1998081698-A/9
PD 31-MAR-1998
PF 24-APR-1997 JP 1997121803
PI 01-MAY-1996 JP 96P 134444
PI TAKEMURA FUJINORI, UENO HIDEKAZU, ITO SATORU
PC C07K1/113,C07K1/14,C12N15/09,C12P21/02,G01N33/531,(C12P21/02,
PC C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH source 1..153 /organism="Mus sp."
FH mat_peptide 1..153 /mol_type="genomic DNA"
FH /db_xref="taxon:10095"
FEATURES
source
1..153
Location/Qualifiers
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
ORIGIN
Query Match 100.0%; Score 12; DB 6; Length 153;
Best Local Similarity 100.0%; Pred.No. 9.7e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
Db 129 TGAGCGCGCGCG 118

RESULT 8
HS37H12R 162 bp DNA linear PRI 23-OCT-1995
LOCUS H.sapiens CpG island DNA genomic MseI fragment, clone 37h12,
DEFINITION reverse read cpg37h12.r1a.
ACCESSION Z65413
VERSION Z65413.1 GI:1038235
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
REFERENCE 2 (bases 1 to 162)
AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT
Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or Contact: biohelp@hgmp.mrc.ac.uk.

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 181)  
McMullen,M.D., Vroh Bi,I., Schroeder,S.S. and Gaut,B.S.  
MPZ-UCI Joint SNP Discovery  
Unpublished (2003)

Contact: Brandon S. Gaut  
Dept. Ecology and Evolutionary Biology  
U.C. Irvine  
321 Steinhaus Hall, Irvine, CA 92697-2525, USA  
Tel: (949) 824-2564  
Fax: (949) 824-2181  
Email: bgaut@uci.edu  
Primer A: gcaaggacaaagatgcgaaaatc  
Primer B: tgcattggcgctataaatgtggg  
STS size: 181

PCR protocol:  
PCR amplification of genomic DNA  
Template: 50 ng  
Primer: each 0.5 uM  
dNTPS: each 200 uM  
Taq Polymerase: Redtaq (Sigma)  
Total Vol: 10 ul

Amplicon sequencing  
ABI protocol - using d-Rhodamine terminator cycle  
sequencing ready reaction with amplitaq DNA polymerase FS  
Sequence ran on ABI 3700 sequencer.

Buffer:  
Genomic DNA amplification  
Redtaq (Sigma)  
Sequencing buffer  
d-Rhodamine kit (ABI)

PHRED/PHRAP Quality Scores 19 20 29 29 19 19 24 29 29 29 29 22 22  
15 20 15 23 8 6 10 9 10 25 22 23 25 23 29 29 40 40 48 48  
34 34 34 25 29 25 27 25 29 37 12 12 12 24 21 18 33 44 42 46 58 69  
69 56 56 55 50 60 64 68 68 71 72 78 76 76 88 81 72 76 76 76 80  
76 81 80 67 62 54 54 54 54 55 59 67 69 77 74 68 59 59 80 64 59  
60 60 63 55 63 64 40 47 47 48 48 54 61 70 86 69 69 71 71  
73 73 88 77 74 68 61 61 59 67 49 39 36 45 47 69 72 65 67 84  
55 61 67 67 65 62 64 66 81 65 73 73 76 69 72 74 72 80 72 81 74  
83 86 77 77 77 77 77 77 80 90 80.

FEATURES  
source  
1..181  
/organism="Zea mays subsp. parviglumis"  
/mol\_type="genomic DNA"  
/cultivar="teos6"  
/db\_xref="taxon:76912"  
/clone\_lib="Zea mays Beadle & Kato site 4"  
/dev\_stage="seedling"  
/note="Organ: leaf; genomic DNA"  
<1..>.181

STS  
ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 181;  
Best Local Similarity 100.0%; Pred. No. 9.4e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGCGGC 12  
| | | | | | | | | |  
Db 13 TGAGCGGCGGC 2

RESULT 11  
BV136872/c  
LOCUS  
DEFINITION PZA00052 Zea mays ssp. parviglumis STS genomic,  
sequence tagged site.  
ACCESSION BV136875.1 GI:45676398  
VERSION BV136875  
KEYWORDS Zea mays subsp. parviglumis  
ORGANISM Zea mays subsp. parviglumis  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

FEATURES  
Location/Qualifiers  
1..162  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="37h12"  
/sex="male"  
/tissue\_type="blood"  
/clone\_lib="CGI-1"  
/dev\_stages="adult"

ORIGIN

Query Match 100.0%; Score 12; DB 9; Length 162;  
Best Local Similarity 100.0%; Pred. No. 9.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGCGGC 12  
| | | | | | | | | |  
Db 102 TGAGCGGCGGC 113

RESULT 9  
HUMSCL1/c  
LOCUS  
DEFINITION Human stem cell leukemia gene product, exon 1a.  
ACCESSION M61103  
VERSION M61103.1 GI:337960  
KEYWORDS stem cell protein.  
SEGMENT 1 of 7  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 177)  
Aplan,P.D., Begley,C.G., Bertness,V., Nussmeier,M., Ezquerro,A.,  
Coligan,J. and Kirsch,I.R.  
The SCL gene is formed from a transcriptionally complex locus  
Mol. Cell. Biol. 10 (12), 6436-6435 (1990)  
91061750  
2247063  
PUBMED  
Original source text: Human DNA.  
Location/Qualifiers  
1..177  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/gene="SCL"  
/gene="SCL"

exon  
intron

ORIGIN

Query Match 100.0%; Score 12; DB 9; Length 177;  
Best Local Similarity 100.0%; Pred. No. 9.5e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGCGGC 12  
| | | | | | | | | |  
Db 121 TGAGCGGCGGC 110

RESULT 10  
BV136875/c  
LOCUS  
DEFINITION PZA00052 Zea mays ssp. parviglumis Beadle & Kato site 4 Zea mays  
Beadle & Kato site 4 Zea mays subsp. parviglumis STS genomic,  
sequence tagged site.  
ACCESSION BV136875  
VERSION BV136875.1 GI:45676398  
KEYWORDS Zea mays subsp. parviglumis  
SOURCE Zea mays subsp. parviglumis  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

KEYWORDS STS.  
 SOURCE Zea mays subsp. parviglumis  
 ORGANISM Zea mays subsp. parviglumis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 183)  
 AUTHORS McMullen,M.D., Vroh Bi,I., Schroeder,S.S. and Gaut,B.S.  
 TITLE MPZ-UCI Joint SNP Discovery  
 JOURNAL Unpublished (2003)  
 COMMENT  
 Contact: Brandon S. Gaut  
 Dept. Ecology and Evolutionary Biology  
 U.C. Irvine  
 321 Steinhaus Hall, Irvine, CA 92697-2525, USA  
 Tel: (949) 824-2564  
 Fax: (949) 824-2181  
 Email: bgaut@uci.edu  
 Primer A: gacaggacagatgcgcaaaatc  
 Primer B: tgcattgcgcttataaatgtgg  
 STS size: 183  
 Protocol:  
 PCR amplification of genomic DNA  
 Template: 50 ng  
 Primer: each 0.5 uM  
 dNTPs: each 200 uM  
 Taq Polymerase: RedTaq (Sigma)  
 Total Vol: 10 ul  
 Amplicon sequencing  
 ABI protocol - using d-Rhodamine terminator cycle  
 sequencing ready reaction with amplitaq DNA polymerase FS  
 Sequence ran on ABI 3700 sequencer.  
 Buffer:  
 Genomic DNA amplification  
 RedTaq (Sigma)  
 Sequencing buffer  
 d-Rhodamine kit (ABI)  
 PHRED/PHRAP Quality Scores 42 42 42 56 42 40 40 38 38 38 31 31 31  
 56 51 37 37 37 27 27 27 42 43 46 46 46 42 42 42 42 42 46  
 37 37 37 51 51 46 65 53 58 56 67 67 73 73 67 66 65 64 65 65 70  
 70 70 75 80 66 66 66 51 46 46 44 26 26 26 34 34 34 51 56 68 69  
 75 66 66 66 75 71 71 58 58 58 60 55 57 57 55 55 55 55 55  
 59 55 55 55 48 55 52 55 47 16 19 24 37 32 32 32 16 16 16  
 28 28 33 52 45 42 34 39 34 32 29 31 27 27 30 30 46 48 87 88  
 88 88 75 70 55 56 47 28 28 29 27 22 18 18 12 13 13 40 28 28  
 28 32 32 22 20 21 22 42 43 43 57 70 64 69 44 41.  
 FEATURES Location/Qualifiers  
 source 1..183  
 /organism="Zea mays subsp. parviglumis"  
 /mol\_type="genomic DNA"  
 /cultivar="teoli"  
 /db\_xref="taxon:76912"  
 /clone\_lib="Zea mays USDA P1566686"  
 /dev\_stage="seedling"  
 /note="Organ: leaf; genomic DNA"  
 <1..183  
 STS  
 ORIGIN  
 Query Match 100.0%; Score 12; DB 11; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TGAGCGGGCGG 12  
 Db 13 TGAGCGGGCGG 2  
 RESULT 12  
 CO657470/c  
 LOCUS 184 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 2396 from Patent WO02070737.  
 Buffer:  
 Genomic DNA amplification  
 RedTaq (Sigma)  
 Sequencing buffer  
 d-Rhodamine kit (ABI)

ACCESSION CO657470  
 VERSION CO657470.1 GI:42120230  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Liew,C.C., Marshall,W.B. and Zhang,H.  
 TITLE Compositions and methods relating to osteoarthritis  
 JOURNAL Patent: WO 02070737-A 2396 12-SEP-2002;  
 Chondrogene Inc. (CA)  
 FEATURES Location/Qualifiers  
 source 1..184  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 100.0%; Score 12; DB 6; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TGAGCGGGCGG 12  
 Db 126 TGAGCGGGCGG 115  
 RESULT 13  
 BV136873/c  
 LOCUS 184 bp DNA linear STS 24-MAR-2004  
 DEFINITION PZA00052 Zea mays ssp. parviglumis JSGYMAS 264 Zea mays JSGYMAS 264  
 Zea mays subsp. parviglumis STS genomic, sequence tagged site.  
 ACCESSION BV136873  
 VERSION BV136873.1 GI:45676396  
 KEYWORDS STS.  
 SOURCE Zea mays subsp. parviglumis  
 ORGANISM Zea mays subsp. parviglumis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 184)  
 AUTHORS McMullen,M.D., Vroh Bi,I., Schroeder,S.S. and Gaut,B.S.  
 TITLE MPZ-UCI Joint SNP Discovery  
 JOURNAL Unpublished (2003)  
 COMMENT  
 Contact: Brandon S. Gaut  
 Dept. Ecology and Evolutionary Biology  
 U.C. Irvine  
 321 Steinhaus Hall, Irvine, CA 92697-2525, USA  
 Tel: (949) 824-2564  
 Fax: (949) 824-2181  
 Email: bgaut@uci.edu  
 Primer A: gacaggacagatgcgcaaaatc  
 Primer B: tgcattgcgcttataaatgtgg  
 STS size: 184  
 Protocol:  
 PCR amplification of genomic DNA  
 Template: 50 ng  
 Primer: each 0.5 uM  
 dNTPs: each 200 uM  
 Taq Polymerase: RedTaq (Sigma)  
 Total Vol: 10 ul  
 Amplicon sequencing  
 ABI protocol - using d-Rhodamine terminator cycle  
 sequencing ready reaction with amplitaq DNA polymerase FS  
 Sequence ran on ABI 3700 sequencer.  
 Buffer:  
 Genomic DNA amplification  
 RedTaq (Sigma)  
 Sequencing buffer  
 d-Rhodamine kit (ABI)

PHRED/PHRAP Quality Scores 66 63 63 59 50 44 47 43 56 56 55 59 59  
 62 62 59 59 63 65 76 78 90 76 73 71 71 74 73 73 69 64 59 59 59 57  
 56 56 55 60 51 56 60 56 56 76 76 82 82 77 77 77 90 88 85 81 78 80  
 61 63 68 70 67 69 79 90 80 83 73 42 42 42 42 42 42 56 56 56 56 48 48  
 48 48 56 56 51 61 61 57 58 61 71 71 71 77 79 76 73 71 71 56 56  
 50 50 47 47 47 56 46 40 40 40 55 57 57 59 59 59 59 59 52 52  
 52 57 62 57 42 42 42 42 42 42 33 33 33 33 42 42 42 42 44 44  
 42 36 36 26 42 38 38 44 56 42 34 34 42 42 42 42 40 40 36 36  
 36 36 35 24 24 44 39 39 39 39 39 39 39 39 42 42.

# FEATURES

Location/Qualifiers  
 1..184  
 /organism="Zea mays subsp. parviglumis"  
 /mol\_type="genomic DNA"  
 /cultivar="teol2"  
 /db\_xref="taxon:76912"  
 /clone\_lib="Zea mays JSCyMAS 264"  
 /dev\_stage="seedling"  
 /note="Organ: leaf; genomic DNA"  
 <1..>184

# STS

## ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGGCG 12

DB 64 TGAGCGGGCGGCG 53

# RESULT 14

BV136865/c  
 LOCUS BV136865 186 bp DNA linear STS 24-MAR-2004  
 DEFINITION PZA00052 Zea mays ssp. mays CML69 Zea mays CML69 Zea mays STS  
 genomic, sequence tagged site.

ACCESSION BV136865

VERSION BV136865.1 GI:45676388

KEYWORDS STS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS McMullen,M.D., Vroh Bi,I., Schroeder,S.S. and Gaut,B.S.

TITLE MPZ-UCI Joint SNP Discovery

JOURNAL Unpublished (2003)

COMMENT

Contact: Brandon S. Gaut  
 Dept. Ecology and Evolutionary Biology  
 U.C. Irvine

321 Steinhaus Hall, Irvine, CA 92697-2525, USA

Tel: (949) 824-2564

Fax: (949) 824-2181

Email: bgaut@uci.edu

Primer A: gacaagacaagatcgcaaaatc

Primer B: tgcattgcgcgtataaatgtggg

STS size: 186

Protocol:

PCR amplification of genomic DNA

Template: 50 ng

Primer: each 0.5 uM

dNTPs: each 200 uM

Taq Polymerase: RedTaq (Sigma)

Total Vol: 10 ul

Amplicon sequencing

ABI protocol - using d-Rhodamine terminator cycle

sequencing ready reaction with amplitaq DNA polymerase FS

Sequence ran on ABI 3700 sequencer.

Buffer:

Genomic DNA amplification

RedTaq (Sigma)  
 Sequencing buffer  
 d-Rhodamine kit (ABI)

PHRED/PHRAP Quality Scores 40 31 31 31 31 31 31 35 35 40 66 64 67 62  
 59 59 59 52 52 56 59 59 55 59 59 61 54 55 42 40 36 51 54 76  
 71 71 71 66 60 60 54 54 59 54 64 54 57 50 56 56 50 43 56 56 56  
 56 56 56 38 38 38 38 38 38 38 38 38 38 38 43 43 51 46 46 37  
 37 37 43 43 56 71 71 57 58 58 66 66 55 58 57 57 57 59 59 57 52  
 52 50 52 37 42 56 48 48 48 56 56 56 56 56 56 44 44 44 44 44  
 40 40 40 40 40 42 32 32 32 32 37 37 29 29 29 29 37 37 42 46 40  
 40 40 40 42 42 32 32 42 42 42 44 37 37 37 37 37 40 40 35 35  
 35 35 35 35 30 30 30 30 30 42 56 56 56 56 41 40 37.

# FEATURES

source

Location/Qualifiers  
 1..186  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="CML69"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Zea mays CML69"  
 /dev\_stage="seedling"  
 /note="Organ: leaf; genomic DNA from inbred line"  
 <1..>186

# STS

## ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 9.4e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGGCG 12

DB 45 TGAGCGGGCGGCG 34

# RESULT 15

BV136855/c

LOCUS BV136855 193 bp DNA linear STS 24-MAR-2004

DEFINITION PZA00052 Zea mays ssp. mays B73(1) Zea mays B73(1) Zea mays STS

genomic, sequence tagged site.

ACCESSION BV136855

VERSION BV136855.1 GI:45676378

KEYWORDS STS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS McMullen,M.D., Vroh Bi,I., Schroeder,S.S. and Gaut,B.S.

TITLE MPZ-UCI Joint SNP Discovery

JOURNAL Unpublished (2003)

COMMENT

Contact: Brandon S. Gaut

Dept. Ecology and Evolutionary Biology

U.C. Irvine

321 Steinhaus Hall, Irvine, CA 92697-2525, USA

Tel: (949) 824-2564

Fax: (949) 824-2181

Email: bgaut@uci.edu

Primer A: gacaagacaagatcgcaaaatc

Primer B: tgcattgcgcgtataaatgtggg

STS size: 193

Protocol:

PCR amplification of genomic DNA

Template: 50 ng

Primer: each 0.5 uM

dNTPs: each 200 uM

Taq Polymerase: RedTaq (Sigma)

Total Vol: 10 ul

Amplicon sequencing

ABI protocol - using d-Rhodamine terminator cycle

sequencing ready reaction with amplitaq DNA polymerase FS

Sequence ran on ABI 3700 sequencer.

```
Buffer:
Genomic DNA amplification
  Redtaq (Sigma)
Sequencing buffer
  d-Rhodamine kit (ABI)

PHRED/PHRAP Quality Scores 36 38 45 44 38 45 45 55 55 55 48 49 48
49 46 44 35 20 19 19 26 33 40 32 32 40 40 40 40 20 20 14
14 17 20 18 29 37 44 40 40 40 37 54 55 64 58 76 74 74 69 73
60 60 60 55 52 54 37 37 35 11 11 11 11 11 45 38 38 56 56
42 42 42 51 51 51 56 56 51 51 51 51 45 40 40 46 35 35 40
37 35 26 26 35 35 40 42 42 42 42 56 51 51 56 56 56 49 49 49
56 40 31 31 31 37 39 39 39 39 39 39 39 39 40 37 37 28 28 28
46 43 43 56 51 35 35 35 35 35 45 40 40 40 56 56 40 45 45 45
51 56 56 45 45 45 40 40 40 40 43 56 56 56 56 56 40 40 40
40 40 35 35.

FEATURES
  source
    1..193
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /cultivar="B73(1)"
      /db_xref="taxon:4577"
      /clone_lib="Zea mays B73(1)"
      /dev_stage="seedling"
      /note="Organ: leaf; genomic DNA from inbred line"
    <1..>193

STS
ORIGIN
  Query Match      100.0%; Score 12; DB 11; Length 193;
  Best Local Similarity 100.0%; Pred. No. 9.3e+04;
  Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGCGCGCG 12
Db      47 TGAGCGCGCGCG 36
```

Search completed: October 28, 2004, 23:44:02  
Job time : 66.1441 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 20:36:09 ; Search time 125.083 Seconds  
(without alignments)  
3495.894 Million cell updates/sec

Title: US-09-867-159A-7

Perfect score: 12

Sequence: 1 tgaaggcgcg 12

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	100.0	60	1	AL588055
2	12	100.0	82	4	BI739668
3	12	100.0	85	8	BH221541
4	12	100.0	100	6	CA389856
5	12	100.0	105	5	BQ044827
6	12	100.0	105	5	BY001184
7	12	100.0	105	7	CN137137
8	12	100.0	106	9	CC689971
9	12	100.0	112	4	BM847049
10	12	100.0	113	6	CA034423
11	12	100.0	116	8	BZ548258
12	12	100.0	116	9	CL193456
13	12	100.0	118	1	AA738907
14	12	100.0	118	1	AA463527
15	12	100.0	118	9	CC684750
16	12	100.0	121	7	CN236655
17	12	100.0	122	4	BI136864
18	12	100.0	122	8	BZ532484
19	12	100.0	127	8	BH221540
20	12	100.0	129	6	CD964482
21	12	100.0	130	4	BI396349
22	12	100.0	133	1	AV436471
23	12	100.0	133	5	BY021939
24	12	100.0	134	4	BG036482

25	12	100.0	134	4	BJ779400
c 26	12	100.0	139	1	AA683894
c 27	12	100.0	139	2	BE241980
28	12	100.0	141	1	AV437091
29	12	100.0	141	1	AV437871
30	12	100.0	141	1	AV438007
31	12	100.0	141	4	BG469858
32	12	100.0	141	4	BG549307
c 33	12	100.0	142	8	BH914993
c 34	12	100.0	142	8	CC200427
35	12	100.0	143	2	AW368325
36	12	100.0	145	2	AW430842
37	12	100.0	145	5	BY062904
c 38	12	100.0	146	6	CD568015
c 39	12	100.0	148	1	AV436751
c 40	12	100.0	150	1	AA064363
41	12	100.0	150	1	AV437772
42	12	100.0	150	1	AV438712
43	12	100.0	150	2	BE359188
44	12	100.0	152	1	AV437504
45	12	100.0	152	1	AV438569

#### ALIGNMENTS

RESULT 1  
AL588055  
LOCUS AL588055 60 bp mRNA linear EST 02-MAR-2001  
DEFINITION AL588055 BP Chicken Brain Library Gallus gallus cDNA clone  
ACCESSION ROS067B07, mRNA sequence.  
VERSION AL588055  
KEYWORDS AL588055.1 GI:13193089  
SOURCE EST.  
ORGANISM Gallus gallus (chicken)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 60)  
Murray, F.  
AUTHORS BP Chicken Brain Library  
TITLE Unpublished (2001)  
JOURNAL  
COMMENT Contact: Frazer Murray  
Dept. Genomics and Bioinformatics  
Roslin, Midlothian, EH25 9PS, UK  
Tel: +44 (0)131 527 4200  
Fax: +44 (0)131 440 0434  
Email: frazer.murray@bbsrc.ac.uk  
GCGGCCGCTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetechn  
(\*6854-  
Seq primer: T7.  
Location/Qualifiers  
1..60  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="ROS067B07"  
/issue\_type="Brain"  
/dev\_stage="Unknown"  
/lab\_host="DH10B"  
/clone\_lib="BP Chicken Brain Library"  
/note="Vector: pSPCRL1; Site 1: Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo dir. 5' adaptor sequence: 5' TCAGACTCGAG 3' ; 3' adaptor sequence: 5' GCGGCCGCTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetechn (\*6854-1)"

#### FEATURES

Source  
1..60  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
/clone="ROS067B07"  
/issue\_type="Brain"  
/dev\_stage="Unknown"  
/lab\_host="DH10B"  
/clone\_lib="BP Chicken Brain Library"  
/note="Vector: pSPCRL1; Site 1: Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo dir. 5' adaptor sequence: 5' TCAGACTCGAG 3' ; 3' adaptor sequence: 5' GCGGCCGCTTTTTTTTTTTTTTTT 3' Poly A RNA purchased from Clonetechn (\*6854-1)"

#### ORIGIN

Query Match 100.0%; Score 12; DB 1; Length 60;  
Best Local Similarity 100.0%; Pred. No. 4e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGG 12  
 Db 34 TGAGCGCGCGG 45

RESULT 2  
 B1739668/c  
 LOCUS  
 DEFINITION 603361935F1 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:5369072 5',  
 mRNA sequence.  
 ACCESSION B1739668  
 VERSION B1739668  
 KEYWORDS B1739668.1 GI:15716681  
 SOURCE EST.  
 ORGANISM Mus musculus (house mouse)

REFERENCE  
 AUTHORS Mus musculus  
 TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 COMMENT NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Inyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLA111939 row: e column: 09  
 High quality sequence stop: 61.  
 Location/Qualifiers

FEATURES  
 source  
 1..82  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5369072"  
 /tissue\_type="retina"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_94"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 3.3 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 100.0%; Score 12; DB 4; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGG 12  
 Db 61 TGAGCGCGCGG 50

RESULT 3  
 BH221541  
 LOCUS  
 DEFINITION 106101G09.y1 1006 - RescueMu Grid G Zea mays genomic, genomic  
 survey sequence.  
 ACCESSION BH221541  
 VERSION BH221541.1 GI:16815800  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 85)  
 Walbot,V.

TITLE  
 JOURNAL  
 COMMENT

Maize genomic sequences found using engineered RescueMu transposon  
 Unpublished (2001)  
 Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu  
 Possible ligation site so sequence was trimmed. Post-ligation  
 sequence submitted separately.  
 Plate: 106101 row: 31  
 Class: transposon-tagged.  
 Location/Qualifiers

FEATURES  
 source  
 1..85  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="mixed background W23/A188/B73"  
 /db\_xref="taxon:4577"  
 /tissue\_type="leaf"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="1006 - RescueMu Grid G"  
 /note="Organ: leaf; Vector: RescueMu (engineered from  
 pBluescript backbone); Site 1: BamHI; Site 2: BglII;  
 RescueMu is a 4.9 kb, modified maize Mu transposon  
 designed to allow plasmid rescue from total genomic DNA.  
 Mu elements insert preferentially into transcription  
 units. For more information on RescueMu, go to the web  
 site 'www.zmdb.iastate.edu' and follow the links for  
 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was  
 extracted from leaf punches, double digested using BamHI  
 and BglII, and ligated to form circular plasmids. DH10B  
 cells were transformed and then screened on LB plates with  
 ampicillin."

ORIGIN  
 Query Match 100.0%; Score 12; DB 8; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGG 12  
 Db 51 TGAGCGCGCGG 62

RESULT 4  
 CA389856  
 LOCUS  
 DEFINITION cs103e01.y1 Human Retinal pigment epithelium/choroid cDNA  
 (Un-normalized, unamplified); cs Homo sapiens cDNA clone cs103e01  
 5', mRNA sequence.  
 ACCESSION CA389856  
 VERSION CA389856.1 GI:24720414  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 100)  
 Wislow,G., Bernstein,S.L., Wyatt,M.K., Farris,R.N., Behal,A.,  
 Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.  
 Expressed sequence tag analysis of human RPE/choroid for the  
 NEIBank Project: Over 6000 non-redundant transcripts, novel genes  
 and splice variants  
 Mol. Vis. 8 (4), 205-220 (2002)

JOURNAL  
 MEDLINE 22103460  
 PUBMED 12107410  
 COMMENT Contact: Wislow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452



/db_xref="taxon:10090"			
/clone="IMAGE:5696219"			
/tissue_type="whole brain"			
/dev_stage="embryo 18.5 dpc"			
/lab_host="PH108 (T1 phage resistant)"			
/clone_lib="NIH BMAP EH0P"			
/note="organ: brain; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."			
ORIGIN			
Query Match		100.0%; Score 12; DB 5; Length 105;	
Best Local Similarity		100.0%; Pred. No. 3.7e-04;	
Matches		12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	TGAGCGGCGCGC 12	
DB	73	TGAGCGGCGCGC 84	
RESULT 6			
BY001184			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
BY001184 105 bp mRNA linear EST 06-DEC-2002			
musculus cDNA clone 1540000GZ1 5', mRNA sequence.			
BY001184			
EST.			
BY001184.1 GI:26061433			
Mus musculus (house mouse)			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 105)			
Okazaki,Y., Osato,N., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojohori,T., Baldarelli,R., Hill,D.P., Bult,C., Batalov,S., Bessel,K.W., Blake,I.L.M., Kanpin,A., Matsuda,H., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.P., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yaginuma,M., Yang,I., Yang,L., Yuan,Z., Zavalan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Akazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Haehizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.			



```

LOCUS       CC689971                106 bp    DNA        linear    GSS 19-JUN-2003
DEFINITION  CGWB35TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM8Ma0531E22,
genomic survey sequence.
ACCESSION   CC689971
VERSION     CC689971.1  GI:32094747
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1  (bases 1 to 106)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Rudiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE       Consortium for Maize Genomics
JOURNAL     Unpublished (2002)
COMMENT     Other GSSs: OGWBT35TV
Contact: Cathy Whitelaw

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES             Location/Qualifiers
     source          1..106
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZM8Ma0531E22"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="vector_PEC5K-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"

ORIGIN
Query Match      100.0%; Score 12; DB 9; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TGAGCGCGCGCG 12
    |||||
DB   69  TGAGCGCGCGCG 58

RESULT 9
BM847049/c
LOCUS       BM847049                112 bp    mRNA        linear    EST 06-MAR-2002
DEFINITION  K-EST0126197 S13KMS5 Homo sapiens cDNA clone S13KMS5-39-G04 5',
mRNA sequence.
ACCESSION   BM847049
VERSION     BM847049.1  GI:19203448
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1  (bases 1 to 112)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongseung@mail.kribb.re.kr
Plate: 39 row: G column: 04

LOCUS       CC689971                106 bp    DNA        linear    GSS 19-JUN-2003
DEFINITION  CGWB35TH_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM8Ma0531E22,
genomic survey sequence.
ACCESSION   CC689971
VERSION     CC689971.1  GI:32094747
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1  (bases 1 to 106)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Rudiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE       Consortium for Maize Genomics
JOURNAL     Unpublished (2002)
COMMENT     Other GSSs: OGWBT35TV
Contact: Cathy Whitelaw

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES             Location/Qualifiers
     source          1..106
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZM8Ma0531E22"
                     /clone_lib="ZM_0.7_1.5_KB"
                     /note="vector_PEC5K-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"

ORIGIN
Query Match      100.0%; Score 12; DB 9; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TGAGCGCGCGCG 12
    |||||
DB   69  TGAGCGCGCGCG 58

RESULT 9
BM847049/c
LOCUS       BM847049                112 bp    mRNA        linear    EST 06-MAR-2002
DEFINITION  K-EST0126197 S13KMS5 Homo sapiens cDNA clone S13KMS5-39-G04 5',
mRNA sequence.
ACCESSION   BM847049
VERSION     BM847049.1  GI:19203448
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1  (bases 1 to 112)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongseung@mail.kribb.re.kr
Plate: 39 row: G column: 04

```

```

High quality sequence stop: 112.
Location/Qualifiers
  1..112
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="S13KMS5-39-G04"
  /tissue_type="myeloma"
  /cell_line="KMS-5"
  /lab_host="Top10F"
  /clone_lib="S13KMS5"
  /note="Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly
  (A)+ RNA was dephosphorylated with bacterial alkaline
  phosphatase (BAP) and then decapped with tobacco acid
  pyrophosphatase (TAP). The decapped intact mRNA was
  ligated with DNA-RNA linker including EcoRI site by
  treatment of T4 RNA ligase and the first strand cDNA was
  synthesized from oligo dt-selected mRNA by priming with
  dt-tailed vector. The dt-tailed vector was adjusted to
  have about 60nt. The cDNA vector was circularized with E.
  coli DNA ligase after digestion of EcoRI which site is
  also included in vector. An RNA strand converted to a DNA
  strand by Okayama-Berg method. The obtained cDNA vectors
  were used for transformation of competent cells E. coli
  Top10F, by electroporation method. The cDNA libraries
  constructed by this method are full-length enriched cDNA
  library."

ORIGIN
Query Match      100.0%; Score 12; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  TGAGCGCGCGCG 12
    |||||
DB   91  TGAGCGCGCGCG 80

RESULT 10
CA034423
LOCUS       CA034423                113 bp    mRNA        linear    EST 24-OCT-2002
DEFINITION  4000046 BARC 5BOV Bos taurus cDNA clone 5BOV_79H05 5', mRNA
sequence.
ACCESSION   CA034423
VERSION     CA034423.1  GI:24332174
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1  (bases 1 to 113)
Sonstegard,T., Capuco,A.V., White,J., Van Tassel,C.P.,
Connor,E.E., Cho,J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D.
and Quackenbush,J.
Analysis of bovine mammary gland EST and functional annotation of
the Bos taurus gene index
Mamm. Genome 13 (7), 373-379 (2002)
12140684
Contact: Sonstegard TS
USDA, ARS, Beltsville Agricultural Research Center
Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
Tel: 301 504 8416
Fax: 301 504 8414
Email: tads@psi.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt ' ' -trim fasta. Vector identified
by cross match using options -mismatch 12 -mismatch 12
Plate: 79 row: H column: 05
Seq primer: ATTGAGTGACACTATAG
High quality sequence stop: 113.
Location/Qualifiers

```

```

source
1..113
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="SBOV 79H05"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="BARC SBOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled mRNA isolated from mammary
tissues at eight physiological, developmental, and disease
states."

ORIGIN
Query Match 100.0%; Score 12; DB 6; Length 113;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
Db 20 TGAGCGCGCGCG 31

RESULT 11
BZ548258
LOCUS
DEFINITION
GAI1Z31TC ZM2_0.7_1.5_KB Zea mays genomic clone ZMMEMa0072E14,
genomic survey sequence.
ACCESSION
BZ548258
VERSION
BZ548258.1 GI:27104919
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 116)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
JOURNAL
Contact: Cathy Whiteaw
COMMENT
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: 1F
Class: sheared ends.
FEATURES
Location/Qualifiers
1..116
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMEMa0072E14"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 12; DB 8; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
Db 85 TGAGCGCGCGCG 96

RESULT 12
CL193456

```

```

LOCUS
DEFINITION
104_417_10940946_114_32274_077 Sorghum methylation-filtered library
(libID: 104) Sorghum bicolor genomic clone 10940946, genomic survey
sequence.
ACCESSION
CL193456
VERSION
CL193456.1 GI:40705979
KEYWORDS
GSS.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 116)
Budiman,M.A., Flick,E., Jones,J., Nunberg,A., Citek,R.W.,
Robbins,D., Rohlfing,T., Bradford,K., Fries,J., McMenamy,J.,
Trani,L., Isak,A., Zimmerman,C., Lakey,N. and Bedell,J.A.
GeneThresher methylation filtered genomic sequences from Sorghum
bicolor
JOURNAL
Unpublished (2004)
COMMENT
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 417 row: c column: 18
Seq primer: M13/pUC Forward
Class: shotgun
High quality sequence stop: 116.
FEATURES
Location/Qualifiers
1..116
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="10940946"
/clone_lib="Sorghum methylation-filtered library (libID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation-filtered library."

ORIGIN
Query Match 100.0%; Score 12; DB 9; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.6e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
Db 92 TGAGCGCGCGCG 103

RESULT 13
AA738907
LOCUS
DEFINITION
v62d12.11 Soares thymus_2NBMT Mus musculus cDNA clone
IMAGE:1226999 5', mRNA sequence.
ACCESSION
AA738907
VERSION
AA738907.1 GI:2776159
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 118)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

```

TITLE  
JOURNAL  
COMMENT

The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouse@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:652591  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 95.  
Location/Qualifiers  
1. .118  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1226999"  
/sex="male"  
/tissue\_type="Thymus"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares thymus 2NBMT"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAATGGAGCGCGCTTATTTTITTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through two  
rounds of normalization, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 12; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGCG 12  
|||||

Db 27 TGAGCGCGCGCG 38  
|||||

RESULT 14  
AA463527

LOCUS  
DEFINITION  
zx72a07.r1 Soares\_total\_fetus Nb2HF8\_9w Homo sapiens cDNA clone  
IMAGE:796980 5', mRNA sequence.

ACCESSION  
AA463527

VERSION  
AA463527.1 GI:2188411

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 118)

AUTHORS  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,  
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,  
White,I., Wylie,T., Waterston,R. and Wilson,R.  
WashU-Merck EST Project 1997  
Unpublished (1997)  
Contact: Willson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the

TITLE  
JOURNAL  
COMMENT

The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouse@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:652591  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 95.  
Location/Qualifiers  
1. .118  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1226999"  
/sex="male"  
/tissue\_type="Thymus"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares thymus 2NBMT"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAATGGAGCGCGCTTATTTTITTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through two  
rounds of normalization, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 12; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGCG 12  
|||||

Db 27 TGAGCGCGCGCG 38  
|||||

RESULT 14  
AA463527

LOCUS  
DEFINITION  
zx72a07.r1 Soares\_total\_fetus Nb2HF8\_9w Homo sapiens cDNA clone  
IMAGE:796980 5', mRNA sequence.

ACCESSION  
AA463527

VERSION  
AA463527.1 GI:2188411

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 118)

AUTHORS  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,  
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,  
White,I., Wylie,T., Waterston,R. and Wilson,R.  
WashU-Merck EST Project 1997  
Unpublished (1997)  
Contact: Willson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the

TITLE  
JOURNAL  
COMMENT

The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouse@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:652591  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 95.  
Location/Qualifiers  
1. .118  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:6041701"  
/db\_xref="taxon:9606"  
/clone="IMAGE:796980"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares total fetus Nb2HF8\_9w"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from pooled 8-9 week  
(total) fetus material with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAATGGAGCGCGCTTATTTTITTTTTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 12; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGCG 12  
|||||

Db 87 TGAGCGCGCGCG 98  
|||||

RESULT 15  
CC684750/c

LOCUS  
DEFINITION  
CGUHL06TV ZM\_0.7.1.5 KB Zea mays genomic clone ZMMBMA0441A12,  
genomic survey sequence.

ACCESSION  
CC684750

VERSION  
CC684750.1 GI:32089526

KEYWORDS  
GSS.

SOURCE  
Zea mays

ORGANISM  
Zea mays

REFERENCE  
1 (bases 1 to 118)

AUTHORS  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)

TITLE  
JOURNAL

COMMENT  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: Tg  
Class: sheared ends.

FEATURES  
source  
1. .118  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0441A12"  
/note="Vector: pSCSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28ml3 rev2 ET from Amersham  
High quality sequence stop: 104.  
Location/Qualifiers  
1. .118  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:6041701"  
/db\_xref="taxon:9606"  
/clone="IMAGE:796980"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares total fetus Nb2HF8\_9w"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from pooled 8-9 week  
(total) fetus material with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAATGGAGCGCGCTTATTTTITTTTTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 12; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGCG 12  
|||||

Db 87 TGAGCGCGCGCG 98  
|||||

RESULT 15  
CC684750

LOCUS  
DEFINITION  
CGUHL06TV ZM\_0.7.1.5 KB Zea mays genomic clone ZMMBMA0441A12,  
genomic survey sequence.

ACCESSION  
CC684750

VERSION  
CC684750.1 GI:32089526

KEYWORDS  
GSS.

SOURCE  
Zea mays

ORGANISM  
Zea mays

REFERENCE  
1 (bases 1 to 118)

AUTHORS  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)

TITLE  
JOURNAL

COMMENT  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: Tg  
Class: sheared ends.

FEATURES  
source  
1. .118  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0441A12"  
/note="Vector: pSCSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

Sun Oct 31 13:10:41 2004

Query Match 100.0%; Score 12; DB 9; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+04;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TGAGCGGGCGG 12  
 |||||  
 Db 77 TGAGCGGGCGG 66

Search completed: October 29, 2004, 01:43:46  
 Job time : 132.083 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: .October 28, 2004, 19:39:58 ; Search time 14.8821 Seconds  
(without alignments)  
4232.808 Million cell updates/sec

Title: US-09-867-159A-7

Perfect score: 12

Sequence: 1 tgagcgcgcgcg 12

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_23Sep04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	100.0	12	6	AAL41283
2	12	100.0	12	6	ABQ80834
3	12	100.0	20	4	AAD14759
4	12	100.0	21	11	ADJ13944
5	12	100.0	21	11	ADJ13804
6	12	100.0	29	9	ACC70267
7	12	100.0	29	9	ACC70269
8	12	100.0	51	3	AAC27652
9	12	100.0	99	2	AAT30923
10	12	100.0	100	2	AAT30927
11	12	100.0	133	2	AAV06342
12	12	100.0	183	6	ABQ91838
13	12	100.0	185	12	ADQ05429
14	12	100.0	186	6	ABQ90654
15	12	100.0	186	6	ABQ91050
16	12	100.0	189	6	ABQ91492
17	12	100.0	189	6	ABQ91274
18	12	100.0	192	6	ABQ90538
19	12	100.0	198	11	ABQ08991
20	12	100.0	216	5	AAD13908
21	12	100.0	219	5	AAH52279

C	22	12	100.0	222	6	ABQ91649
C	23	12	100.0	236	3	AAC08685
C	24	12	100.0	285	6	ABQ91398
C	25	12	100.0	285	10	ABX82177
C	26	12	100.0	286	10	ABX88122
C	27	12	100.0	291	6	ABQ91575
C	28	12	100.0	300	10	ABZ39199
C	29	12	100.0	301	3	AAC10194
C	30	12	100.0	306	3	AAC01927
C	31	12	100.0	309	9	ADA14495
C	32	12	100.0	316	6	ABN15903
C	33	12	100.0	322	8	ABZ55310
C	34	12	100.0	324	6	ABQ91632
C	35	12	100.0	327	3	AAC07364
C	36	12	100.0	333	3	AA31815
C	37	12	100.0	342	10	ADB80120
C	38	12	100.0	348	6	ABL86085
C	39	12	100.0	366	9	ACH20867
C	40	12	100.0	373	8	ABX52363
C	41	12	100.0	378	8	ACA26590
C	42	12	100.0	387	11	ABD17392
C	43	12	100.0	396	5	AA64828
C	44	12	100.0	408	6	ABK76658
C	45	12	100.0	410	8	ABX51978

## ALIGNMENTS

RESULT 1  
AAL41283  
ID AAL41283 standard; DNA; 12 BP.  
XX  
AC AAL41283;  
XX  
DT 02-JAN-2003 (first entry)  
XX  
DE Oligonucleotide of DP/DF epitope.  
XX  
KW Antiallergic; antiinflammatory; antiasthmatic; dermatological; allergen;  
KW anti-histamine; histamine synthesis inhibitor; allergic hypersensitivity;  
KW allergic asthma; allergic rhinitis; atopic eczema; DP/DF; epitope; ds.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
PN WO200278736-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-FR001098.  
XX  
PR 30-MAR-2001; 2001FR-00004370.  
PR 03-MAY-2001; 2001FR-00005929.  
PR 29-MAY-2001; 2001US-00867159.  
PA (ANTI-) ANTIALIS SARK.  
XX  
PI Loria E, Terrasse G, Trehin Y;  
XX  
WPI; 2002-750636/81.  
XX  
PT Antiallergic compositions containing an anti-histamine, a histamine  
FT synthesis inhibitor, and optionally an allergen or nucleic acid coding  
FT for the allergen.  
XX  
PS Claim 20; Page 13; 32pp; French.  
XX  
CC The invention relates to antiallergic compositions containing an anti-  
CC histamine, a histamine synthesis inhibitor, and optionally an allergen or  
CC isolated nucleic acid molecule that has at least one polynucleotide  
CC sequence coding for the allergen, together with a pharmaceutical carrier.  
CC The pharmaceutical composition of the invention is useful as a non-  
CC specific antiallergic treatment, and also useful in the treatment of

CC allergic hypersensitivity, allergic asthma, allergic rhinitis, and  
CC allergic and atopic eczema. This polynucleotide sequence represents an  
CC oligonucleotide of a DP/DF epitope relating to the anti-allergic  
CC compositions of the invention  
XX  
SQ Sequence 12 BP; 1 A; 3 C; 7 G; 1 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 12; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.5e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TGAGCGGCGCG 12  
Db 1 TGAGCGGCGCG 12  
  
RESULT 2  
ABQ80834  
ID ABQ80834 standard; DNA; 12 BP.  
XX  
AC ABQ80834;  
XX  
DT 13-DEC-2002 (first entry)  
XX  
DE Cysteine protease PCR primer.  
XX  
KW Antiallergic; antiasthmatic; antiinflammatory; dermatological;  
KW immunotherapy; allergen; allergic hypersensitivity reaction;  
KW allergic asthma; allergic rhinitis; allergic atopic eczema;  
KW cysteine protease; PCR; primer; ss.  
XX  
OS Dermatophagoides pteronyssinus.  
XX  
PN FR222709-A1.  
XX  
PD 04-OCT-2002.  
XX  
PF 03-MAY-2001; 2001FR-00005929.  
XX  
PR 30-MAR-2001; 2001FR-00004370.  
XX  
PA (ANTI-) ANTIALIS SARL.  
XX  
PI Loria E, Terrasse G, Trehin Y;  
XX  
DR WPI; 2002-735037/80.  
XX  
PT Antiallergic composition, useful for preventing and treating e.g. asthma,  
PT rhinitis or eczema, containing at least two of allergen, antihistamine  
PT and histamine synthesis inhibitor.  
XX  
PS Claim 19; Page 7; 33pp; French.  
XX  
CC The present invention relates to an anti-allergic pharmaceutical  
CC composition (I) comprising a pharmaceutical carrier containing an active  
CC agent combination of at least two of: an allergen; an antihistamine; and  
CC a histamine synthesis inhibitor. (I) is used for treating or preventing  
CC allergic hypersensitivity reactions, especially allergic asthma, allergic  
CC rhinitis or allergic atopic eczema, in babies, children or adults. To  
CC illustrate the invention, cysteine protease from Dermatophagoides  
CC pteronyssinus (AB98533), was used as an allergen. The present sequence  
CC is a PCR primer, used to illustrate the invention  
XX  
SQ Sequence 12 BP; 1 A; 3 C; 7 G; 1 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 12; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.5e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 TGAGCGGCGCG 12  
Db 1 TGAGCGGCGCG 12

RESULT 3  
AAD14759/c  
ID AAD14759 standard; DNA; 20 BP.  
XX  
AC AAD14759;  
XX  
DT 01-NOV-2001 (first entry)  
XX  
DE Human glycogen synthase kinase 3 alpha antisense oligo ISIS #116600.  
XX  
KW Human; glycogen synthase kinase 3 alpha; antidiabetic; cytostatic;  
KW antisense therapy; diabetes; hyperproliferative disorder; inflammation;  
KW neurological disorder; tumour; haematopoietic disorder; infection;  
KW hyperproliferative disorder; developmental disorder; antisense;  
KW phosphorothioate backbone; ss.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PH Key Location/Qualifiers  
FT modified\_base 1..20  
FT /tag= a  
FT /mod\_base= OTHER  
FT /note= "Phosphorothioate backbone"  
FT modified\_base 1..5  
FT /tag= b  
FT /mod\_base= OTHER  
FT /note= "Methoxyethyl residues"  
FT modified\_base 5  
FT /tag= d  
FT /mod\_base= m5c  
FT modified\_base 6  
FT /tag= e  
FT /mod\_base= m5c  
FT modified\_base 7  
FT /tag= f  
FT /mod\_base= m5c  
FT modified\_base 9  
FT /tag= g  
FT /mod\_base= m5c  
FT modified\_base 10  
FT /tag= h  
FT /mod\_base= m5c  
FT modified\_base 12  
FT /tag= i  
FT /mod\_base= m5c  
FT modified\_base 13  
FT /tag= j  
FT /mod\_base= m5c  
FT modified\_base 15  
FT /tag= k  
FT /mod\_base= m5c  
FT modified\_base 16..20  
FT /tag= c  
FT /mod\_base= OTHER  
FT modified\_base 17  
FT /note= "Methoxyethyl residues"  
FT /tag= l  
FT /mod\_base= m5c  
XX  
PN WC200152865-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 16-JAN-2001; 2001WO-US001411.  
XX  
PR 21-JAN-2000; 2000US-00488856.  
XX  
PA (ISIS-) ISIS PHARM INC.  
XX  
PI Monia BP, McKay R, Butler MM, Wyatt JR;  
XX



```

DR WPI; 2001-442247/47.
XX
PT Antisense compound 8 to 30 nucleobases in length comprising a compound
PT that is targeted to a nucleic acid molecule encoding glycogen synthase
PT kinase 3 alpha, useful for the treatment of e.g. diabetes and
PT hyperproliferative disorders.
XX
XX Example 15; Page 82; 115pp; English.
XX
CC The invention relates to an antisense compound 8 to 30 nucleobases in
CC length targeted to a nucleic acid encoding glycogen synthase kinase 3
CC alpha. The antisense compound specifically hybridises with and inhibits
CC the expression of glycogen synthase kinase 3 alpha. The antisense
CC compound is useful for the treatment of a diseases associated with
CC glycogen synthase kinase 3 alpha such as diabetes, a neurological
CC disorder, a haematopoietic disorder, a hyperproliferative disorder or a
CC developmental disorder. The antisense compounds may also be used
CC prophylactically to prevent or delay infection, inflammation or tumour
CC formation. The present sequence is a phosphorothioate antisense
CC oligonucleotide targetted to human glycogen synthase kinase 3 alpha DNA
XX
SQ Sequence 20 BP; 3 A; 9 C; 6 G; 2 T; 0 U; 0 Other;
    Query Match      100.0%; Score 12; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 9.1e+03;
    Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    XX
    QY 1 TGAGCGCGCGCG 12
    Db 18 TGAGCGCGCGCG 7
    |||||
    |||||

RESULT 4
ADJ13944/c
ID ADJ13944 standard; DNA; 21 BP.
XX
AC ADJ13944;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human DNA probe used to immobilise CpG methylated DNA SeqID 1071.
XX
DE probe; ss; chemical modification; methylation; array; CpG island;
XX tumour suppressor; p16; human; H69; H1618.
XX
OS Homo sapiens.
XX
PN US2003152950-A1.
XX
PD 14-AUG-2003.
XX
PF 27-JUN-2002; 2002US-00184085.
XX
PR 27-JUN-2001; 2001US-0301370P.
XX
XX (GARN/) GARNER H R.
XX (MINN/) MINNA J D.
XX (LUEB/) LUEBKE K J.
XX (BALO/) BALOG R P.
XX
PI Garner HR, Minna JD, Luebke KJ, Balog RP;
XX
DR WPI; 2003-874843/81.
XX
XX Analysis of chemical modification of DNA involves obtaining sample of DNA
XX to be analyzed, treating DNA with chemical reagents that result in
XX PT different base sequences, and determining sequence of resulting DNA.
XX
XX Example 1; SEQ ID NO 1071; 210pp; English.
XX
XX This invention relates to a novel method for analysing chemically
XX modified macromolecules. Specifically, it refers to a high throughput
XX method for the parallel analysis of many potential sites of chemical
XX modification (e.g. methylation) in DNA. The present invention describes
XX treating the DNA with one or more chemical reagents that result in
XX different base sequences depending upon the presence or absence of the
XX modification of interest. Accordingly, a device comprising an array of
XX probes is provided to hybridise with and select the altered DNA sequences
XX that comprise the modifications of interest such as a CpG island. In
XX particular, this invention refers to analysing the methylation pattern of
XX a region of the promoter for the tumour suppressor gene p16 from two
XX human lung tumour cell lines H69 and H1618. This oligonucleotide sequence
XX

WPI; 2003-874843/81.
XX
XX Analysis of chemical modification of DNA involves obtaining sample of DNA
XX to be analyzed, treating DNA with chemical reagents that result in
XX PT different base sequences, and determining sequence of resulting DNA.
XX
XX Example 1; SEQ ID NO 931; 210pp; English.
XX
XX This invention relates to a novel method for analysing chemically
XX modified macromolecules. Specifically, it refers to a high throughput
XX method for the parallel analysis of many potential sites of chemical
XX modification (e.g. methylation) in DNA. The present invention describes
XX treating the DNA with one or more chemical reagents that result in
XX different base sequences depending upon the presence or absence of the
XX modification of interest. Accordingly, a device comprising an array of
XX probes is provided to hybridise with and select the altered DNA sequences
XX that comprise the modifications of interest such as a CpG island. In
XX particular, this invention refers to analysing the methylation pattern of
XX a region of the promoter for the tumour suppressor gene p16 from two
XX human lung tumour cell lines H69 and H1618. This oligonucleotide sequence
XX

```

CC is a human DNA probe used to immobilise CpG methylated DNA of the  
CC invention.

XX Sequence 21 BP; 3 A; 11 C; 4 G; 3 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 12; DB 11; Length 21;

XX Best Local Similarity 100.0%; Pred. No. 9.1e+03;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12

DB 15 TGAGCGCGCGCG 4

## RESULT 6

ACC70267/c  
ID ACC70267 standard; DNA; 29 BP.

XX AC ACC70267;

XX 11-AUG-2003 (first entry)

XX Primer for DNA encoding EHV-1 full length glycoprotein gg.

XX Chemokine; glycoprotein; gg; alphaherpesvirus infection;

XX infectious disease; malaria; tuberculosis; HIV; inflammatory disease;

XX psoriasis; atherosclerosis; neoplasia; diabetes; colitis; BHV-5;

XX transplant rejection; PCR; primer; ss.

XX Bovine herpesvirus 5.

XX WO2003038440-A2.

XX 08-MAY-2003.

XX 30-OCT-2002; 2002WO-GB004918.

XX 30-OCT-2001; 2001GB-00026047.

XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX (ANIM-) ANIMAL HEALTH TRUST.

XX Alcamì A, Bryant N, Davis-Poynter N;

XX WPI; 2003-468433/44.

XX Obtaining a modulator of chemokine activity comprises bringing into  
XX contact a glycoprotein (gg), a chemokine polypeptide and a test compound,  
XX and determining interaction or binding between the gg and chemokine  
XX polypeptides.

XX Disclosure; Page 64; 86pp; English.

XX The specification describes a method of obtaining a modulator of  
XX chemokine activity. The method comprises bringing into contact

XX glycoprotein gg, chemokine polypeptides and a test compound, under

XX conditions where gg and chemokine polypeptides bind in the absence of the

XX test compound being an inhibitor; and determining interaction or binding

XX between the gg and chemokine polypeptides. The gg polypeptides, or

XX polynucleotides and antibodies are useful for the manufacture of a

XX medicament for treating alphaherpesvirus infection or a disorder

XX involving chemokine activity e.g., infectious diseases (malaria,

XX tuberculosis or HIV), inflammatory diseases, psoriasis, atherosclerosis,

XX neoplasia, diabetes, colitis or transplant rejection. PCR primers

XX ACC70267-68 were used to amplify DNA encoding a viral gg. The amplified

XX product was cloned and expressed in a baculovirus system, for use in the

XX method of the invention

XX Sequence 29 BP; 5 A; 13 C; 7 G; 4 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 100.0%; Score 12; DB 9; Length 29;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12

DB 26 TGAGCGCGCGCG 15

## RESULT 7

ACC70269/c

ID ACC70269 standard; DNA; 29 BP.

XX AC ACC70269;

XX 11-AUG-2003 (first entry)

XX Primer for DNA encoding EHV-1 secreted glycoprotein gg.

XX Chemokine; glycoprotein; gg; alphaherpesvirus infection;

XX infectious disease; malaria; tuberculosis; HIV; inflammatory disease;

XX psoriasis; atherosclerosis; neoplasia; diabetes; colitis; BHV-5;

XX transplant rejection; PCR; primer; ss.

XX Bovine herpesvirus 5.

XX WO2003038440-A2.

XX 08-MAY-2003.

XX 30-OCT-2002; 2002WO-GB004918.

XX 30-OCT-2001; 2001GB-00026047.

XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX (ANIM-) ANIMAL HEALTH TRUST.

XX Alcamì A, Bryant N, Davis-Poynter N;

XX WPI; 2003-468433/44.

XX Obtaining a modulator of chemokine activity comprises bringing into  
XX contact a glycoprotein (gg), a chemokine polypeptide and a test compound,  
XX and determining interaction or binding between the gg and chemokine  
XX polypeptides.

XX Disclosure; Page 64; 86pp; English.

XX The specification describes a method of obtaining a modulator of  
XX chemokine activity. The method comprises bringing into contact  
XX glycoprotein gg, chemokine polypeptides and a test compound, under

XX conditions where gg and chemokine polypeptides bind in the absence of the

XX test compound being an inhibitor; and determining interaction or binding

XX between the gg and chemokine polypeptides. The gg polypeptides, or

XX polynucleotides and antibodies are useful for the manufacture of a

XX medicament for treating alphaherpesvirus infection or a disorder

XX involving chemokine activity e.g., infectious diseases (malaria,

XX tuberculosis or HIV), inflammatory diseases, psoriasis, atherosclerosis,

XX neoplasia, diabetes, colitis or transplant rejection. PCR primers

XX ACC70269-70 were used to amplify DNA encoding a viral gg. The amplified

XX product was cloned and expressed in a baculovirus system, for use in the

XX method of the invention

XX Sequence 29 BP; 5 A; 13 C; 7 G; 4 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 100.0%; Score 12; DB 9; Length 29;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12

DB 26 TGAGCGCGCGCG 15

## RESULT 8

AAC27652/c

ID AC227652 standard; cDNA; 51 BP.  
XX AAC27652;  
AC  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein 5' EST, SEQ ID NO: 31727.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-00200610.  
XX PR 26-FEB-1999; 99US-0122487P.  
XX PA (GEST) GENSET.  
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors  
XX SQ Sequence 51 BP; 14 A; 18 C; 11 G; 7 T; 0 U; 1 Other;  
  
Query Match 100.0%; Score 12; DB 3; Length 51;  
Best Local Similarity 100.0%; Pred. No. 8.4e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TGAGCGGGCGG 12  
Db |||||  
33 TGAGCGGGCGG 22  
  
RESULT 9  
AAT30923  
ID AAT30923 standard; DNA; 99 BP.  
XX  
XX AC AAT30923;  
XX DT 24-FEB-1997 (first entry)  
XX DE Primer F11 for 80 kD protein coding sequence.  
XX KW Tetrahymena; 80 kD protein; 95 kD protein; telomerase; telomeric repeat;  
KW RNA coupled protein; chromosome; telomere repeat synthesis; antibody;  
KW immortalised cell; cancer; eukaryotic microbe; inhibitor; gene therapy;  
KW fungal infection; therapy; diagnosis; protozoan infection; somatic cell;

KW p80; p95; polymerase chain reaction; primer; amplify; PCR; ss.  
XX Synthetic.  
OS  
XX PN WO9619580-A2.  
XX  
XX PD 27-JUN-1996.  
XX PF 18-DEC-1995; 95WO-US016531.  
XX PR 19-DEC-1994; 94US-00359125.  
XX PA (COLD-) COLD SPRING HARBOR LAB.  
XX Greider C, Collins K, Kobayashi R, Yang XH, Hemish JM;  
PI Autexier C;  
XX WPI; 1996-309594/31.  
XX PT Telomerase protein and related DNA, antibodies, transgenic cells, etc. -  
PT for diagnosis and treatment of cancer and infection by eukaryotic  
PT microbes, also new telomerase inhibitors.  
XX Claim 36; Fig 9; 56pp; English.  
XX AAT30871-730931 represent amplification primers for the coding sequences  
CC for the Tetrahymena 80 kD and 95 kD proteins (p80 and p95), respectively  
CC (see AAT30867 and AAT30868 for wild type coding sequences). AAT30905-  
CC T3031 represent amplification primers for p80. The proteins encoded by  
CC the amplified sequences are components of telomerase. The encoded  
CC proteins act, when coupled to RNA, to add telomeric repeats (of the  
CC sequence TTGGGG) to stabilise chromosomes. The RNA component provides the  
CC template for the telomere repeat synthesis. Antibodies against the  
CC encoded proteins can be used as immunoassay reagents for the detection  
CC of the proteins. The antibodies can also be used to identify immortalised  
CC cells, or predisposition to immortalisation, particularly cancer, or to  
CC diagnose a disease caused by a eukaryotic microbe. Inhibitors of the  
CC encoded proteins (and the proteins themselves) can be used for therapy or  
CC diagnosis. The inhibitors can be used to treat infection by fungi and  
CC protozoa. As somatic cells do not generally require telomerase, these  
CC inhibitors should have little or no toxicity to the host. The amplified  
CC sequences are used to produce the recombinant protein, or to isolate  
CC similar genes from other organisms, while transformed cells can be used  
CC in gene therapy  
XX SQ Sequence 99 BP; 20 A; 20 C; 36 G; 23 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 12; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 7.9e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TGAGCGGGCGG 12  
Db |||||  
3 TGAGCGGGCGG 14  
  
RESULT 10  
AAT30927/c  
ID AAT30927 standard; DNA; 100 BP.  
XX  
XX AC AAT30927;  
XX DT 24-FEB-1997 (first entry)  
XX DE Primer R10 for 80 kD protein coding sequence.  
XX KW Tetrahymena; 80 kD protein; 95 kD protein; telomerase; telomeric repeat;  
KW RNA coupled protein; chromosome; telomere repeat synthesis; antibody;  
KW immortalised cell; cancer; eukaryotic microbe; inhibitor; gene therapy;  
KW fungal infection; therapy; diagnosis; protozoan infection; somatic cell;  
KW p80; p95; polymerase chain reaction; primer; amplify; PCR; ss.  
OS Synthetic.

```

XX WO9619580-A2.
PN 27-JUN-1996.
XX 18-DEC-1995; 95WO-US016531.
PD 19-DEC-1994; 94US-00359125.
XX (COLD-) COLD SPRING HARBOR LAB.
XX Greider C, Collins K, Kobayashi R, Yang XH, Hemish JM;
PI Autexier C;
XX WPI; 1996-309594/31.
XX Telomerase protein and related DNA, antibodies, transgenic cells, etc. -
PT for diagnosis and treatment of cancer and infection by eukaryotic
PT microbes, also new telomerase inhibitors.
XX Claim 36; Fig 9; 56pp; English.
XX AAT30871-T30931 represent amplification primers for the coding sequences
CC for the Tetrahymena 80 kD and 95 kD proteins (p80 and p95), respectively
CC (see AAT30867 and AAT30868 for wild type coding sequences). AAT30905-
CC T30931 represent amplification primers for p80. The proteins encoded by
CC the amplified sequences are components of telomerase. The encoded
CC proteins act, when coupled to RNA, to add telomeric repeats (of the
CC sequence TTGGGG) to stabilise chromosomes. The RNA component provides the
CC template for the telomere repeat synthesis. Antibodies against the
CC encoded proteins, can be used as immunoassay reagents for the detection
CC of the proteins. The antibodies can also be used to identify immortalised
CC cells, or predisposition to immortalisation, particularly cancer, or to
CC diagnose a disease caused by a eukaryotic microbe. Inhibitors of the
CC encoded proteins (and the proteins themselves) can be used for therapy or
CC diagnosis. The inhibitors can be used to treat infection by fungi and
CC protozoa. As somatic cells do not generally require telomerase, these
CC inhibitors should have little or no toxicity to the host. The amplified
CC sequences are used to produce the recombinant protein, or to isolate
CC similar genes from other organisms, while transformed cells can be used
CC in gene therapy
XX Sequence 100 BP; 13 A; 31 C; 29 G; 27 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 12; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.9e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
DB 17 TGAGCGCGCGCG 6

RESULT 11
AAV06342/C
ID AAV06342 standard; DNA; 153 BP.
XX AC AAV06342;
XX 30-APR-1998 (first entry)
XX Mouse protamine 1 encoding DNA.
XX Nucleic acid-binding motif; HBc; protamine 1; mouse; diagnosis;
KW HCV core polypeptide; immunoassay; detection; antigen; disease;
KW Hepatitis C virus; infection; ss.
XX Synthetic.
OS Mus sp.
XX Location/Qualifiers
FH Key 1..153
FT CDS /*tag= a

```

```

FT /product= "Mouse protamine 1"
FT /note= "the stop codon is not indicated"
XX EP805160-A1.
XX 05-NOV-1997.
XX 30-APR-1997; 97EP-00406985.
XX 01-MAY-1996; 96JP-00134444.
XX (FJRE ) FUJIREBIO INC.
XX Takemura F, Ueno E, Itoh S;
XX WPI; 1997-529030/49.
XX P-PSDB; AAW33694.
XX Nucleic acid-bound polypeptide - useful as immunoassay reagent.
XX Example 15; Page 32; 38pp; English.
XX This DNA encodes a mouse protamine 1. This is used in a method for
CC producing a nucleic acid-bound polypeptide. The method comprises
CC producing a fusion gene containing the polypeptide and a nucleic acid-
CC binding motif, binding a nucleic acid to the polypeptide as a soluble
CC fraction, and purifying the nucleic acid-bound polypeptide from the
CC soluble fraction. When the polypeptide is a recombinant form of an
CC antigen, the nucleic acid-bound polypeptide can be used as an immunoassay
CC reagent for detecting the antigen or an antibody to the antigen,
CC especially in an agglutination assay using particles coated with the
CC nucleic acid-bound polypeptide. The methods can be applied to diagnosis
CC of disease and infection, especially for the detection of HBV and HCV
CC polypeptides. The nucleic acid-bound polypeptides may be immunoreactive
CC in cases where the free polypeptide is not
XX Sequence 153 BP; 42 A; 41 C; 51 G; 19 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 12; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 7.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12
DB 129 TGAGCGCGCGCG 118

RESULT 12
ABQ91838/C
ID ABQ91838 standard; DNA; 183 BP.
XX AC ABQ91838;
XX 01-OCT-2002 (first entry)
XX M. capsulatus gene #1823 for DNA array.
XX Micro array; gene; ds; differential expression; gene expression.
XX Methylococcus capsulatus.
XX WO200255655-A2.
XX 18-JUL-2002.
XX 14-JAN-2002; 2002WO-NC0000019.
XX 12-JAN-2001; 2001NO-00000235.
XX 12-JAN-2001; 2001NO-00000239.
XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
PA (TIGR-) TIGR.
XX

```

PI Birkeland NK, Bidhammer I, Jonassen I, Jensen HB, Lien T;  
PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;  
PI Salzberg SL;  
XX WPI; 2002-557818/59.  
XX Novel DNA array useful for determining differential expression of  
PT Methylococcus capsulatus genes, comprises polynucleotides or  
PT oligonucleotides representative for a selective number of Methylococcus  
PT capsulatus genes.  
XX Claim 14; Page 673; 678pp; English.  
XX The invention relates to a novel DNA array giving a representation of a  
CC number of Methylococcus capsulatus genes. The method of the invention is  
CC useful for determination of the differential expression of the genes of  
CC M. capsulatus, and for studying gene expression on a genomic scale and in  
CC gene expression assays of M. capsulatus genes. The sequences shown in  
CC ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the  
XX invention  
XX Sequence 183 BP; 27 A; 54 C; 55 G; 47 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 12; DB 6; Length 183;  
Best Local Similarity 100.0%; Pred. NO. 7.5e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TGAGCGGCGGCG 12  
DB 111 TGAGCGGCGGCG 100  
RESULT 13  
ADQ05429/c  
ID ADQ05429 standard; DNA; 185 BP.  
AC ADQ05429;  
XX 09-SEP-2004 (first entry)  
XX Soybean leucine zipper transcription factor seqid 2275.  
XX maize; soybean; Arabidopsis thaliana; transcription factor; homeobox;  
KW HLH; leucine zipper; zinc finger; transformed plant; plant;  
KW metabolic pathway; mutation detection; polymorphism; plant trait;  
KW genome mapping; gene identification; gene analysis; plant breeding;  
transgenic; ds.  
XX Glycine max.  
OS  
XX US2004123339-A1.  
FN  
XX 24-JUN-2004.  
PD  
XX 06-AUG-2001; 2001US-00922293.  
PF  
XX 24-NOV-1997; 97US-0067000P.  
PR 09-DEC-1997; 97US-0069472P.  
PR 13-JAN-1998; 98US-0071479P.  
PR 10-FEB-1998; 98US-0074201P.  
PR 10-FEB-1998; 98US-0074280P.  
PR 10-FEB-1998; 98US-0074281P.  
PR 10-FEB-1998; 98US-0074282P.  
PR 12-FEB-1998; 98US-0074585P.  
PR 12-FEB-1998; 98US-0074586P.  
PR 12-FEB-1998; 98US-0074567P.  
PR 19-FEB-1998; 98US-0074789P.  
PR 19-FEB-1998; 98US-0075459P.  
PR 19-FEB-1998; 98US-0075460P.  
PR 19-FEB-1998; 98US-0075461P.  
PR 19-FEB-1998; 98US-0075462P.  
PR 19-FEB-1998; 98US-0075463P.  
PR 19-FEB-1998; 98US-0075464P.

PR 09-MAR-1998; 98US-0077229P.  
PR 09-MAR-1998; 98US-0077230P.  
PR 09-MAR-1998; 98US-0077231P.  
PR 18-MAR-1998; 98US-0078368P.  
PR 07-APR-1998; 98US-0080844P.  
PR 27-APR-1998; 98US-0083067P.  
PR 29-APR-1998; 98US-0083386P.  
PR 29-APR-1998; 98US-0083387P.  
PR 29-APR-1998; 98US-0083388P.  
PR 29-APR-1998; 98US-0083389P.  
PR 13-MAY-1998; 98US-0085222P.  
PR 13-MAY-1998; 98US-0085223P.  
PR 13-MAY-1998; 98US-0085224P.  
PR 15-MAY-1998; 98US-0085533P.  
PR 21-MAY-1998; 98US-0086183P.  
PR 21-MAY-1998; 98US-0086184P.  
PR 21-MAY-1998; 98US-0086185P.  
PR 21-MAY-1998; 98US-0086186P.  
PR 21-MAY-1998; 98US-0086187P.  
PR 21-MAY-1998; 98US-0086188P.  
PR 16-JUN-1998; 98US-0089524P.  
PR 18-JUN-1998; 98US-0089793P.  
PR 18-JUN-1998; 98US-0089806P.  
PR 18-JUN-1998; 98US-0089807P.  
PR 18-JUN-1998; 98US-0089808P.  
PR 18-JUN-1998; 98US-0089810P.  
PR 18-JUN-1998; 98US-0089811P.  
PR 18-JUN-1998; 98US-0089812P.  
PR 18-JUN-1998; 98US-0089813P.  
PR 18-JUN-1998; 98US-0089814P.  
PR 30-JUN-1998; 98US-0091247P.  
PR 30-JUN-1998; 98US-0091405P.  
PR 09-SEP-1998; 98US-0099667P.  
PR 09-SEP-1998; 98US-0099668P.  
PR 09-SEP-1998; 98US-0099670P.  
PR 09-SEP-1998; 98US-0099672P.  
PR 16-SEP-1998; 98US-0100672P.  
PR 16-SEP-1998; 98US-0100673P.  
PR 16-SEP-1998; 98US-0100674P.  
PR 17-SEP-1998; 98US-0100963P.  
PR 21-SEP-1998; 98US-0101130P.  
PR 21-SEP-1998; 98US-0101131P.  
PR 21-SEP-1998; 98US-0101132P.  
PR 22-SEP-1998; 98US-0101343P.  
PR 22-SEP-1998; 98US-0101344P.  
PR 22-SEP-1998; 98US-0101347P.  
PR 22-SEP-1998; 98US-0101508P.  
PR 25-SEP-1998; 98US-0101707P.  
PR 13-OCT-1998; 98US-0104124P.  
PR 13-OCT-1998; 98US-0104125P.  
PR 13-OCT-1998; 98US-0104127P.  
PR 13-OCT-1998; 98US-0104128P.  
PR 18-NOV-1998; 98US-0108966P.  
PR 19-NOV-1998; 98US-0109018P.  
PR 24-NOV-1998; 98US-00199129.  
PR 08-DEC-1998; 98US-00210297.  
PR 11-DEC-1998; 98US-0111981P.  
PR 22-DEC-1998; 98US-0113224P.  
PR 12-JAN-1999; 99US-00229413.  
XX (CONN/) CONNER T W.  
PA (HECK/) HECK G R.  
PA (LIU/) LIU J.  
XX Conner TW, Heck GR, Liu J;  
XX WPI; 2004-468202/44.  
XX New substantially purified nucleic acid molecule that encodes a maize,  
PT soybean or Arabidopsis thaliana transcription factor or its fragment,  
PT useful for genome mapping, gene identification and analysis or plant  
PT breeding.

PS Claim 2; SEQ ID NO 2275; 140pp; English.

XX The invention describes a substantially purified nucleic acid molecule

CC that encodes a maize, soybean or Arabidopsis thaliana transcription

CC factor or its fragment, where the maize or soybean transcription factor

CC is homeobox, HbH, leucine zipper, zinc finger, or other transcription

CC factor. Also described are: a substantially purified maize or soybean

CC transcription factor or its fragment defined above; a substantially

CC purified antibody or its fragment which is capable of specifically

CC binding to the transcription factor or its fragment above; a transformed

CC plant; a method for determining a level or pattern in a plant cell of a

CC transcription factor in a plant metabolic pathway; a method of

CC determining a mutation in a plant; a method of producing a plant

CC containing an overexpressed protein or reduced levels of plant

CC transcription factor; a method of determining an association between a

CC polymorphism and a plant trait; and a method of isolating a nucleic acid

CC that encodes a plant transcription factor or its fragment. The nucleic

CC acid molecules, proteins and their fragments are useful for genome

CC mapping, gene identification and analysis, plant breeding, preparation of

CC constructs for use in plant gene expression and transgenic plants. The

CC nucleic acid molecules are useful as markers or probes. This sequence

CC represents a soybean leucine zipper transcription factor polynucleotide.

XX

SQ Sequence 185 BP; 41 A; 54 C; 47 G; 43 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 12; Length 185;

Best Local Similarity 100.0%; Pred. No. 7.5e+03;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGCG 12

DB 129 TGAGCGGGCGCG 118

RESULT 14

ABQ90654/c

ID ABQ90654 standard; DNA; 186 BP.

AC ABQ90654;

XX

DT 01-OCT-2002 (first entry)

DE M. capsulatus gene #639 for DNA array.

XX

XX Micro array; gene; ds; differential expression; gene expression.

XX

OS Methylococcus capsulatus.

XX

PN WO200255655-A2.

XX

PD 18-JUL-2002.

XX

PF 14-JAN-2002; 2002WO-NO000019.

XX

PR 12-JAN-2001; 2001NO-00000235.

XX

PR 12-JAN-2001; 2001NO-00000239.

XX

XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.

PA (TIGR-) TIGR.

XX

PI Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;

PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;

PI Salzberg SL;

XX

XX WPI; 2002-557818/59.

XX

XX Novel DNA array useful for determining differential expression of

PT Methylococcus capsulatus genes, comprises polynucleotides or

PT oligonucleotides representative for a selective number of Methylococcus

XX capsulatus genes.

PS Claim 14; Page 316; 678pp; English.

XX

XX The invention relates to a novel DNA array giving a representation of a

XX number of Methylococcus capsulatus genes. The method of the invention is

XX useful for determination of the differential expression of the genes of

XX M. capsulatus, and for studying gene expression on a genomic scale and in

XX gene expression assays of M. capsulatus genes. The sequences shown in

XX ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the

XX invention

SQ Sequence 186 BP; 29 A; 52 C; 55 G; 50 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 6; Length 186;

Best Local Similarity 100.0%; Pred. No. 7.4e+03;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGCG 12

CC The invention relates to a novel DNA array giving a representation of a

CC number of Methylococcus capsulatus genes. The method of the invention is

CC useful for determination of the differential expression of the genes of

CC M. capsulatus, and for studying gene expression on a genomic scale and in

CC gene expression assays of M. capsulatus genes. The sequences shown in

CC ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the

CC invention

XX

SQ Sequence 186 BP; 32 A; 51 C; 57 G; 46 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 6; Length 186;

Best Local Similarity 100.0%; Pred. No. 7.4e+03;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGCG 12

DB 111 TGAGCGGGCGCG 100

RESULT 15

ABQ91050/c

ID ABQ91050 standard; DNA; 186 BP.

XX

XX AC ABQ91050;

XX

DT 01-OCT-2002 (first entry)

XX

DE M. capsulatus gene #1035 for DNA array.

XX

XX Micro array; gene; ds; differential expression; gene expression.

XX

OS Methylococcus capsulatus.

XX

PN WO200255655-A2.

XX

PD 18-JUL-2002.

XX

XX 14-JAN-2002; 2002WO-NO000019.

XX

PR 12-JAN-2001; 2001NO-00000235.

XX

PR 12-JAN-2001; 2001NO-00000239.

XX

XX (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.

PA (TIGR-) TIGR.

XX

PI Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;

PI Lillehaug JR, Lossius I, Eisen JA, Fraser CM, Durkin AS;

PI Salzberg SL;

XX

XX WPI; 2002-557818/59.

XX

XX Novel DNA array useful for determining differential expression of

PT Methylococcus capsulatus genes, comprises polynucleotides or

PT oligonucleotides representative for a selective number of Methylococcus

XX capsulatus genes.

PS Claim 14; Page 452-453; 678pp; English.

XX

XX The invention relates to a novel DNA array giving a representation of a

XX number of Methylococcus capsulatus genes. The method of the invention is

XX useful for determination of the differential expression of the genes of

XX M. capsulatus, and for studying gene expression on a genomic scale and in

XX gene expression assays of M. capsulatus genes. The sequences shown in

XX ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the

XX invention

SQ Sequence 186 BP; 29 A; 52 C; 55 G; 50 T; 0 U; 0 Other;

Query Match 100.0%; Score 12; DB 6; Length 186;

Best Local Similarity 100.0%; Pred. No. 7.4e+03;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGCG 12

Db 111 TGAGCGGCGG 100

Search completed: October 28, 2004, 22:43:39  
Job time : 19.8821 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 23:44:08 ; Search time 110.079 Seconds  
(without alignments)  
558.975 Million cell updates/sec

Title: US-09-867-159A-7

Perfect score: 12

Sequence: 1 TGAGCGCGGCG 12

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	100.0	12	10	US-09-867-159A-7
2	12	100.0	20	15	US-10-181-875-13
3	12	100.0	21	15	US-10-184-088A-931
4	12	100.0	21	15	US-10-184-088A-1071
5	12	100.0	39	18	US-10-411-910A-291
6	12	100.0	39	18	US-10-411-910A-307
7	12	100.0	39	18	US-10-411-910A-325
8	12	100.0	153	9	US-09-306-780-17
9	12	100.0	165	17	US-10-437-963-36390
10	12	100.0	184	16	US-10-242-538A-2396
11	12	100.0	184	16	US-10-085-783A-2396
12	12	100.0	185	11	US-09-922-293-2275

13	12	100.0	198	17	US-10-437-963-67499	Sequence 67499, A
14	12	100.0	201	17	US-10-741-601-22502	Sequence 22502, A
15	12	100.0	201	17	US-10-741-601-22506	Sequence 22506, A
16	12	100.0	201	17	US-10-741-601-22562	Sequence 22562, A
17	12	100.0	216	16	US-10-182-327-25	Sequence 25, Appl
c 18	12	100.0	225	17	US-10-437-963-22569	Sequence 22569, A
c 19	12	100.0	249	17	US-10-437-963-80657	Sequence 80657, A
c 20	12	100.0	250	17	US-10-437-963-63014	Sequence 63014, A
c 21	12	100.0	250	17	US-10-437-963-83375	Sequence 83375, A
22	12	100.0	252	17	US-10-437-963-34708	Sequence 34708, A
23	12	100.0	264	17	US-10-437-963-76400	Sequence 76400, A
24	12	100.0	268	16	US-10-062-674-1016	Sequence 1016, Ap
25	12	100.0	295	16	US-10-424-599-18255	Sequence 18255, A
26	12	100.0	295	17	US-10-437-963-37924	Sequence 37924, A
27	12	100.0	306	17	US-10-437-963-80916	Sequence 80916, A
c 28	12	100.0	309	17	US-10-437-963-83021	Sequence 83021, A
c 29	12	100.0	310	17	US-10-767-701-18665	Sequence 18665, A
c 30	12	100.0	316	17	US-10-437-963-10995	Sequence 10995, A
c 31	12	100.0	318	17	US-10-437-963-13762	Sequence 13762, A
32	12	100.0	327	17	US-10-437-963-40641	Sequence 40641, A
c 33	12	100.0	333	14	US-10-062-727-776	Sequence 776, App
c 34	12	100.0	333	17	US-10-437-963-86781	Sequence 86781, A
c 35	12	100.0	341	16	US-10-424-599-75779	Sequence 75779, A
c 36	12	100.0	342	17	US-10-437-963-9193	Sequence 9193, Ap
c 37	12	100.0	342	17	US-10-437-963-65058	Sequence 65058, A
c 38	12	100.0	345	15	US-10-156-761-5486	Sequence 5486, Ap
c 39	12	100.0	348	9	US-09-867-701-9063	Sequence 9063, Ap
c 40	12	100.0	348	17	US-10-437-963-93778	Sequence 93778, A
c 41	12	100.0	351	15	US-10-156-761-3083	Sequence 3083, Ap
42	12	100.0	351	17	US-10-437-963-64612	Sequence 64612, A
43	12	100.0	362	13	US-10-027-632-127994	Sequence 127994, A
44	12	100.0	362	15	US-10-027-632-127994	Sequence 127994, A
c 45	12	100.0	364	15	US-10-424-599-24018	Sequence 24018, A

ALIGNMENTS

RESULT 1  
US-09-867-159A-7  
; Sequence 7, Application US/09867159A  
; Publication No. US20030104013A1  
; GENERAL INFORMATION:  
; APPLICANT: ANTIALIS TERRASSE, GAETAN LORIA, EMILE TREHIN, YVES  
; TITLE OF INVENTION: Anti-allergic pharmaceutical composition containing at least one anti-histamine compound  
; FILE REFERENCE: B112812US-antialis  
; CURRENT APPLICATION NUMBER: US/09/867,159A  
; CURRENT FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: FR01/04370  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: FR01/05929  
; PRIOR FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 12  
; TYPE: DNA  
; ORGANISM: Dermatophagoides pteronyssinus  
; FEATURE:  
; NAME/KEY: primer  
; LOCATION: (1)..(12)  
; OTHER INFORMATION:  
US-09-867-159A-7

Query Match 100.0%; Score 12; DB 10; Length 12;  
Best Local Similarity 100.0%; Pred. No. 5.3e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGGCG 12

DB 1 TGAGCGCGGCG 12

```

; APPLICANT: Minna, John D.
; APPLICANT: Luebke, Kevin, J.
; APPLICANT: Balog, Robert P.
; TITLE OF INVENTION: Identification of Chemically Modified Polymers
; FILE REFERENCE: 11929-1035
; CURRENT APPLICATION NUMBER: US/10/184,085A
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/301,370
; PRIOR FILING DATE: 2001-06-27
; NUMBER OF SEQ ID NOS: 1291
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1071
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-184-085A-1071

Query Match      100.0%; Score 12; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGCGCGCG 12
Db      15 TGAGCGCGCGCG 4

RESULT 5
US-10-411-910A-291
; Sequence 291, Application US/10411910A
; Publication No. US20040209256A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Harrison F.
; TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
; FILE REFERENCE: H2041203-P
; CURRENT APPLICATION NUMBER: US/10/411,910A
; CURRENT FILING DATE: 2003-04-12
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 291
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; US-10-411-910A-291

Query Match      100.0%; Score 12; DB 18; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.9e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGCGCGCG 12
Db      8 TGAGCGCGCGCG 19

RESULT 6
US-10-411-910A-307
; Sequence 307, Application US/10411910A
; Publication No. US20040209256A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Harrison F.
; TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
; FILE REFERENCE: H2041203-P
; CURRENT APPLICATION NUMBER: US/10/411,910A
; CURRENT FILING DATE: 2003-04-12
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 307
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; US-10-184-085A-931/c
; Sequence 931, Application US/10184085A
; Publication No. US20030152950A1
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Minna, John D.
; APPLICANT: Luebke, Kevin, J.
; APPLICANT: Balog, Robert P.
; TITLE OF INVENTION: Identification of Chemically Modified Polymers
; FILE REFERENCE: 11929-1035
; CURRENT APPLICATION NUMBER: US/10/184,085A
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/301,370
; PRIOR FILING DATE: 2001-06-27
; NUMBER OF SEQ ID NOS: 1291
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 931
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-184-085A-931

Query Match      100.0%; Score 12; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGCGCGCG 12
Db      18 TGAGCGCGCGCG 7

RESULT 3
US-10-184-085A-931/c
; Sequence 931, Application US/10184085A
; Publication No. US20030152950A1
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Minna, John D.
; APPLICANT: Luebke, Kevin, J.
; APPLICANT: Balog, Robert P.
; TITLE OF INVENTION: Identification of Chemically Modified Polymers
; FILE REFERENCE: 11929-1035
; CURRENT APPLICATION NUMBER: US/10/184,085A
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 60/301,370
; PRIOR FILING DATE: 2001-06-27
; NUMBER OF SEQ ID NOS: 1291
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 931
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-184-085A-931

Query Match      100.0%; Score 12; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGCGCGCG 12
Db      15 TGAGCGCGCGCG 4

RESULT 4
US-10-184-085A-1071/c
; Sequence 1071, Application US/10184085A
; Publication No. US20030152950A1
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
```

US-10-411-910A-307

Query Match 100.0%; Score 12; DB 18; Length 39;  
Best Local Similarity 100.0%; Pred. No. 3.9e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12  
|||||  
DB 8 TGAGCGCGCGCG 19

## RESULT 7

US-10-411-910A-325  
; Sequence 325, Application US/10411910A  
; Publication No. US20040209256A1  
; GENERAL INFORMATION:  
; APPLICANT: Dillon, Harrison F.  
; TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes  
; FILE REFERENCE: H2041203-P  
; CURRENT APPLICATION NUMBER: US/10/411,910A  
; CURRENT FILING DATE: 2003-04-12  
; NUMBER OF SEQ ID NOS: 343  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 325  
; LENGTH: 39  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-411-910A-325

Query Match 100.0%; Score 12; DB 18; Length 39;  
Best Local Similarity 100.0%; Pred. No. 3.9e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12  
|||||  
DB 8 TGAGCGCGCGCG 19

## RESULT 8

US-09-306-780-17/c  
; Sequence 17, Application US/09306780  
; Publication No. US20010051336A1  
; GENERAL INFORMATION:  
; APPLICANT: TAKEMURA, FUMINORI  
; UENO, HIICHI  
; ITOH, SATORU  
; TITLE OF INVENTION: NUCLEIC ACID-BOUND POLYPEPTIDE, METHOD  
; OF PRODUCING NUCLEIC ACID-BOUND POLYPEPTIDE AND  
; IMMUNOASSAY USING THE POLYPEPTIDE.  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/306,780  
; FILING DATE: 07-May-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/841,657A  
; FILING DATE: 30-APR-1997  
; APPLICATION NUMBER: JP 8-134444

FILING DATE: 01-MAY-1997

ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 2084-033-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 153 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "synthetic DNA"

FEATURE:

NAME/KEY: CDS  
; LOCATION: 1..153  
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-306-780-17

Query Match 100.0%; Score 12; DB 9; Length 153;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12  
|||||  
DB 129 TGAGCGCGCGCG 118

## RESULT 9

US-10-437-963-36390  
; Sequence 36390, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 36390  
; LENGTH: 165  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_40220C.1  
US-10-437-963-36390

Query Match 100.0%; Score 12; DB 17; Length 165;  
Best Local Similarity 100.0%; Pred. No. 2.8e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12  
|||||  
DB 78 TGAGCGCGCGCG 89

## RESULT 10

US-10-242-535A-2396/c  
; Sequence 2396, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.

```

; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2396
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-2396

Query Match      100.0%; Score 12; DB 16; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.7e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCAGCGCGCGG 12
DB      126 TCAGCGCGCGG 115

RESULT 11
US-10-085-783A-2396/c
; Sequence 2396, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2396
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-2396

Query Match      100.0%; Score 12; DB 16; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.7e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCAGCGCGCGG 12
DB      126 TCAGCGCGCGG 115

RESULT 12
US-09-922-293-2275/c
; Sequence 2275, Application US/09922293
; Publication No. US20040123339A1
; GENERAL INFORMATION:
; APPLICANT: Conner, Timothy W.
; APPLICANT: Heck, Gregory R.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
```

```

; TITLE OF INVENTION: Transcription in Plants
; FILE REFERENCE: 16517.254
; CURRENT APPLICATION NUMBER: US/09/922,293
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/067,000
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: US 60/069,472
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: US 60/071,479
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/074,201
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,282
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,280
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,281
; PRIOR FILING DATE: 1998-02-10
; PRIOR APPLICATION NUMBER: US 60/074,566
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/074,567
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/074,565
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: US 60/075,462
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/074,789
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,459
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,461
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,464
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,460
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/075,463
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 60/077,231
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 60/077,229
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 60/077,230
; PRIOR FILING DATE: 1998-03-09
; PRIOR APPLICATION NUMBER: US 60/078,368
; PRIOR FILING DATE: 1998-03-18
; PRIOR APPLICATION NUMBER: US 60/080,844
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 60/083,067
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: US 60/083,386
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/083,387
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/083,388
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/083,389
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: US 60/085,224
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,223
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,222
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,533
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/086,186
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,187
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,185
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/086,184
```

; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/086,183  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/086,188  
; PRIOR FILING DATE: 1998-05-21  
; PRIOR APPLICATION NUMBER: US 60/089,524  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: US 60/089,810  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,814  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,808  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,812  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,807  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,806  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,813  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,811  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/089,793  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: US 60/091,405  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/091,247  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/099,667  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 60/099,668  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 60/099,670  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 60/099,697  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 60/100,674  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: US 60/100,673  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: US 60/100,672  
; PRIOR FILING DATE: 1998-09-16  
; PRIOR APPLICATION NUMBER: US 60/100,963  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: US 60/101,131  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US 60/101,132  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US 60/101,130  
; PRIOR FILING DATE: 1998-09-21  
; PRIOR APPLICATION NUMBER: US 60/101,508  
; PRIOR FILING DATE: 1998-09-22  
; PRIOR APPLICATION NUMBER: US 60/101,344  
; PRIOR FILING DATE: 1998-09-22  
; PRIOR APPLICATION NUMBER: US 60/101,347  
; PRIOR FILING DATE: 1998-09-22  
; PRIOR APPLICATION NUMBER: US 60/101,343  
; PRIOR FILING DATE: 1998-09-22  
; PRIOR APPLICATION NUMBER: US 60/101,707  
; PRIOR FILING DATE: 1998-09-25  
; PRIOR APPLICATION NUMBER: US 60/104,126  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: US 60/104,128  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: US 60/104,127  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: US 60/104,124  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: US 60/109,018  
; PRIOR FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: US 60/108,996  
; PRIOR FILING DATE: 1998-11-18

; PRIOR APPLICATION NUMBER: US 09/199,129  
; PRIOR FILING DATE: 1998-11-24  
; PRIOR APPLICATION NUMBER: US 09/210,297  
; PRIOR FILING DATE: 1998-12-08  
; PRIOR APPLICATION NUMBER: US 60/111,981  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: US 60/113,224  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: US 09/229,413  
; PRIOR FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 3853  
; SEQ ID NO 2275  
; LENGTH: 185  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-922-293-2275

Query Match 100.0%; Score 12; DB 11; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12  
DB 129 TGAGCGGGCGG 118

## RESULT 13

US-10-437-963-67499  
; Sequence 67499, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 67499  
; LENGTH: 198  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_68350C.1  
US-10-437-963-67499

Query Match 100.0%; Score 12; DB 17; Length 198;  
Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12  
DB 171 TGAGCGGGCGG 182

## RESULT 14

US-10-741-601-22502  
; Sequence 22502, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: C1001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22

Sun Oct 31 13:10:40 2004

```
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22502
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-22502
```

```
Query Match      100.0%; Score 12; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 TGAGCGCGCGCG 12
          |||||
Db      17 TGAGCGCGCGCG 28
```

```
RESULT 15
US-10-741-601-22506
; Sequence 22506, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22506
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-22506
```

```
Query Match      100.0%; Score 12; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 TGAGCGCGCGCG 12
          |||||
Db      73 TGAGCGCGCGCG 84
```

```
Search completed: October 29, 2004, 06:26:15
Job time : 112.079 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 28, 2004, 22:29:43 ; Search time 3.28384 Seconds  
(without alignments)  
2597.404 Million cell updates/sec

Title: US-09-867-159A-7

Perfect score: 12

Sequence: 1 TGAGCGCGGCG 12

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCRUS COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	12	100.0	20	3	US-09-488-856A-13
C 2	12	100.0	51	4	US-09-513-999C-31727
C 3	12	100.0	178	4	US-09-270-767-30807
C 4	12	100.0	198	4	US-09-252-991A-7595
C 5	12	100.0	236	4	US-09-513-999C-12760
C 6	12	100.0	285	4	US-09-313-294A-637
C 7	12	100.0	286	4	US-09-313-294A-6582
C 8	12	100.0	301	3	US-08-651-155B-143
C 9	12	100.0	301	4	US-09-194-036B-143
C 10	12	100.0	301	4	US-09-513-999C-14269
C 11	12	100.0	306	4	US-09-513-999C-1925
C 12	12	100.0	327	4	US-09-513-999C-11439
C 13	12	100.0	350	4	US-09-270-767-6247
C 14	12	100.0	350	4	US-09-270-767-21529
C 15	12	100.0	387	4	US-09-252-991A-15996
C 16	12	100.0	419	5	PCT-US95-08293-20
C 17	12	100.0	423	4	US-09-252-991A-12661
C 18	12	100.0	426	2	US-08-387-942C-29
C 19	12	100.0	426	4	US-09-397-787-218
C 20	12	100.0	430	4	US-09-621-976-11346
C 21	12	100.0	430	4	US-09-489-039A-5981
C 22	12	100.0	464	4	US-09-621-976-15301
C 23	12	100.0	473	4	US-09-270-767-14602
C 24	12	100.0	480	4	US-09-513-999C-1846
C 25	12	100.0	481	4	US-09-270-767-7745
C 26	12	100.0	481	4	US-09-270-767-23027
C 27	12	100.0	495	4	US-09-252-991A-5255

28 12 100.0 496 1 US-07-847-743B-22 Sequence 22, Appl  
29 12 100.0 496 1 US-08-456-201-22 Sequence 22, Appl  
30 12 100.0 496 2 US-08-456-241-22 Sequence 22, Appl  
31 12 100.0 496 5 PCT-US92-04295A-22 Sequence 22, Appl  
32 12 100.0 504 1 US-07-661-610C-15 Sequence 15, Appl  
33 12 100.0 522 4 US-09-252-991A-8866 Sequence 8866, Ap  
C 34 12 100.0 540 4 US-09-252-991A-4274 Sequence 4274, Ap  
C 35 12 100.0 602 4 US-09-513-999C-14868 Sequence 14868, A  
C 36 12 100.0 622 3 US-08-986-837-1 Sequence 1, Appl  
C 37 12 100.0 622 4 US-09-809-592-1 Sequence 1, Appl  
C 38 12 100.0 654 4 US-09-252-991A-4481 Sequence 4481, Ap  
C 39 12 100.0 685 2 US-08-254-354-1 Sequence 1, Appl  
C 40 12 100.0 685 2 US-08-254-354-3 Sequence 3, Appl  
C 41 12 100.0 685 5 PCT-US95-06137-1 Sequence 1, Appl  
C 42 12 100.0 685 5 PCT-US95-06137-3 Sequence 3, Appl  
C 43 12 100.0 732 4 US-09-489-039A-1297 Sequence 1297, Ap  
C 44 12 100.0 744 4 US-09-252-991A-13301 Sequence 13301, A  
C 45 12 100.0 753 4 US-09-489-039A-6109 Sequence 6109, Ap

#### ALIGNMENTS

RESULT 1  
US-09-488-856A-13/c  
; Sequence 13, Application US/09488856A  
; Patent No. 6316259  
; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia  
; APPLICANT: Robert McKay  
; APPLICANT: Madeline M. Butler  
; APPLICANT: Jacqueline Wyatt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP  
; FILE REFERENCE: RTS-0115  
; CURRENT APPLICATION NUMBER: US/09/488.856A  
; CURRENT FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 88  
; SEQ ID NO 13  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Antisense Oligonucleotide  
US-09-488-856A-13

Query Match 100.0%; Score 12; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGGCG 12

Db 18 TGAGCGCGGCG 7

#### RESULT 2

US-09-513-999C-31727/c  
; Sequence 31727, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 31727  
; LENGTH: 51

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2
; OTHER INFORMATION: v=a or c or g
US-09-513-999C-31727

Query Match      100.0%; Score 12; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGGCGGCG 12
DB      33 TGAGCGGCGGCG 22

RESULT 3
US-09-270-767-30807/c
; Sequence 30807; Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 30807
; LENGTH: 178
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30807

Query Match      100.0%; Score 12; DB 4; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGGCGGCG 12
DB      35 TGAGCGGCGGCG 24

RESULT 4
US-09-252-991A-7595/c
; Sequence 7595; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7595
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7595

Query Match      100.0%; Score 12; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGGCGGCG 12
DB      43 TGAGCGGCGGCG 32

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2
; OTHER INFORMATION: v=a or c or g
US-09-513-999C-12760

Query Match      100.0%; Score 12; DB 4; Length 236;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGGCGGCG 12
DB      118 TGAGCGGCGGCG 129

RESULT 5
US-09-513-999C-12760
; Sequence 12760; Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 12760
; LENGTH: 236
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 74
; OTHER INFORMATION: y=c or t
US-09-513-999C-12760

Query Match      100.0%; Score 12; DB 4; Length 236;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGGCGGCG 12
DB      118 TGAGCGGCGGCG 129

RESULT 6
US-09-313-294A-637
; Sequence 637; Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: halgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 637
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700549568H1
; NAME/KEY: unsure
; LOCATION: 199-222
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-637

Query Match      100.0%; Score 12; DB 4; Length 285;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGAGCGGCGGCG 12
DB      55 TGAGCGGCGGCG 66
```



## RESULT 7

US-09-313-294A-6582  
 ; Sequence 6582, Application US/09313294A  
 ; Patent No. 6476212  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lalgudi, Raghunath V.  
 ; APPLICANT: Ito, Laura Y.  
 ; APPLICANT: Sherman, Bradley K.  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
 ; FILE REFERENCE: PL-0017 US  
 ; CURRENT APPLICATION NUMBER: US/09/313,294A  
 ; CURRENT FILING DATE: 1999-05-14  
 ; NUMBER OF SEQ ID NOS: 7600  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 6582  
 ; LENGTH: 286  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. 6476212 700352053H1  
 ; NAME/KEY: unsure  
 ; LOCATION: 284-285  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-09-313-294A-6582

Query Match 100.0%; Score 12; DB 4; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12

Db 30 TGAGCGCGCGCG 41

## RESULT 8

US-08-651-155B-143/c  
 ; Sequence 143, Application US/08651155B  
 ; Patent No. 6365401  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mahan Dr., Michael J.  
 ; APPLICANT: Conner Mr., Christopher P.  
 ; APPLICANT: Hiethoff Mr., Douglas M.  
 ; TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION  
 ; TITLE OF INVENTION: OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST  
 ; TITLE OF INVENTION: INFECTION  
 ; NUMBER OF SEQUENCES: 255  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Chrisman, Bynum & Johnson, P.C.  
 ; STREET: 1900 Fifteenth Street  
 ; CITY: Boulder  
 ; STATE: CO  
 ; COUNTRY: USA  
 ; ZIP: 80302  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/651,155B  
 ; FILING DATE: 17-MAY-1996  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Petersen Mr., Steven C.  
 ; REGISTRATION NUMBER: 36,238  
 ; REFERENCE/DOCKET NUMBER: 17060.1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 303/546-1300  
 ; TELEFAX: 303/448-5426  
 ; TELEX: ABAL75  
 ; INFORMATION FOR SEQ ID NO: 143:  
 ; SEQUENCE CHARACTERISTICS:

Query Match 100.0%; Score 12; DB 4; Length 301;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;

; LENGTH: 301 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: YES  
 US-08-651-155B-143

Query Match 100.0%; Score 12; DB 3; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGCGCGCG 12

Db 198 TGAGCGCGCGCG 187

## RESULT 9

US-09-194-036B-143/c  
 ; Sequence 143, Application US/09194036B  
 ; Patent No. 6548246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mahan, Michael J.  
 ; Conner, Christopher P.  
 ; Hiethoff, Douglas M.  
 ; TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION  
 ; TITLE OF INVENTION: OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST  
 ; TITLE OF INVENTION: INFECTION  
 ; NUMBER OF SEQUENCES: 255  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morrison & Foerster  
 ; STREET: 755 Page Mill Road  
 ; CITY: Mountain View  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/194,036B  
 ; FILING DATE: 17-No. 6548246-1998  
 ; CLASSIFICATION: <unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US97/08208  
 ; FILING DATE: 1997-05-16  
 ; APPLICATION NUMBER: US 08/651,155  
 ; FILING DATE: 1996-05-17  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Shantanu Basu  
 ; REGISTRATION NUMBER: 43,318  
 ; REFERENCE/DOCKET NUMBER: 220002060601  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 813-5995  
 ; TELEFAX: (650) 494-0792  
 ; INFORMATION FOR SEQ ID NO: 143:  
 ; SEQUENCE CHARACTERISTICS:

Query Match

Best Local Similarity

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12  
DB 198 TGAGCGGGCGG 187

## RESULT 10

US-09-513-999C-14269  
; Sequence 14269, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59 US2 REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14269  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 236  
; OTHER INFORMATION: y=c or t  
US-09-513-999C-14269

Query Match 100.0%; Score 12; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12  
DB 151 TGAGCGGGCGG 162

## RESULT 11

US-09-513-999C-1925  
; Sequence 1925, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59 US2 REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1925  
; LENGTH: 306  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 61..306  
US-09-513-999C-1925

Query Match 100.0%; Score 12; DB 4; Length 306;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12  
DB 12 TGAGCGGGCGG 23

## RESULT 12

US-09-513-999C-11439  
; Sequence 11439, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59 US2 REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 11439  
; LENGTH: 327  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 140  
; OTHER INFORMATION: r=a or g  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 147  
; OTHER INFORMATION: r=a or g  
; NAME/KEY: misc\_feature  
; LOCATION: 229  
; OTHER INFORMATION: r=a or g  
US-09-513-999C-11439

Query Match 100.0%; Score 12; DB 4; Length 327;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12  
DB 293 TGAGCGGGCGG 304

## RESULT 13

US-09-270-767-6247/c  
; Sequence 6247, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 6247  
; LENGTH: 350  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-6247

Query Match 100.0%; Score 12; DB 4; Length 350;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGAGCGGGCGG 12

Db 203 TGAGCGCGCGC 192

## RESULT 14

US-09-270-767-21529/c  
; Sequence 21529, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21529  
; LENGTH: 350  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-21529

Query Match 100.0%; Score 12; DB 4; Length 350;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGC 12

|||||

Db 203 TGAGCGCGCGC 192

## RESULT 15

US-09-252-991A-15996/c  
; Sequence 15996, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 15996  
; LENGTH: 387  
; TYPE: DNA  
; ORGANISM: *Pseudomonas aeruginosa*  
US-09-252-991A-15996

Query Match 100.0%; Score 12; DB 4; Length 387;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGCGCGCGC 12

|||||

Db 126 TGAGCGCGCGC 115

Search completed: October 29, 2004, 01:48:25  
Job time : 6.28384 secs

